Universität Stuttgart
Geodätisches Institut

# Estimation of atmospheric signals from daily gravity field solution

Bachelorarbeit im Studiengang
**Geodäsie und Geoinformatik**
an der Universität Stuttgart

Dennis Frederic Mattes

Stuttgart, April 2017

**Betreuer:** Prof. Dr.-Ing. Nico Sneeuw
Universität Stuttgart

PD Dr.-Ing. habil. Johannes Engels
Universität Stuttgart

M.Sc. Wei Liu
Universität Stuttgart

# Erklärung der Urheberschaft

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt und auch noch nicht veröffentlicht.

Ort, Datum Unterschrift

In this thesis the question will be solved, if it is possible to extract gravity signals from the atmosphere in the measured gravity field of the earth. At the beginning, this will be done by comparing the Daily Wiese Solutions, which are simulated measured gravity fields of the earth and the atmosphere, to have an idea in which way the data is created and how the basic data have already similarities. In the following analysis, two possibilities are tested to estimate the atmosphere. The analysis is based on the idea that the estimation of the atmosphere proves the atmospheric signals in the gravity field of the earth. One possibility is to use *masks* to extract the desired signals, while the other way is to add the differences of two daily solutions to a start field. There it can be seen, that the first method offers the best correlation with the given atmosphere. The latter method has a lower correlation but has the advantage to require a lower number of atmospheric models. Also the requirement of low pass filters in further atmospheric estimations is shown. Two anomalies which are detected are intensity peaks in the estimated atmosphere, which could be related to extreme weather at that time and also the big variation between points in the upper part of the field and in the lower part.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

## 1.1 Introduction

The atmosphere is heavy, the whole atmosphere has a mass of some $5 \cdot 10^{18}$ kg [1]. There are different ways to measure the gravity field of the earth which is correlated to mass variations from changes on the surface of the earth. This can be used to monitor the melting of icebergs in the polar region which have also a big mass. However, if the atmosphere has such a big mass on the earth, the question occurs, if the atmosphere can be also monitored, by using the measured gravity fields of the earth.

This question becomes more important, when we keep in mind the fact, that the ESA (European Space Agency) is planning a new geodetic satellite mission. One possible mission architecture includes the novel property that it contains one pair of satellites with a polar orbit and a second pair, which has an inclination from around 76 degree. This offers the great possibility that there is a good coverage of the earth and of the polar regions at the same time[2]. If we now regard the new satellite mission, which will measure the gravity field of the earth better than ever before, we have the advantage that we have gravity data from the whole earth within a short processing time period and, if it is possible to extract atmospheric gravity signals out of it, we have at the same time an atmospheric model of the whole earth, representing the atmospheric pressure. This mission offers new possibilities for the atmospheric science, because it offers first hand data, about how the pressure of the atmosphere varies during the year. If we have a look at the current global warming, and take into account that the atmospheric pressure depends on the temperature, it becomes clear that the change in the temperature also affects a change in the atmospheric pressure. Thus, it is more important than before, to monitor also the changes in the atmosphere which can be caused by climatic changes. Therefore, with the possibility to estimate the atmosphere from the gravity data of the earth, a new way of atmospheric monitoring is created.

In this thesis, we will investigate the question, in which way the estimate the time variable impact on the atmospheric gravity field of the gravity field of the earth. To do so, we first have to look at the data we use, even here before we really start with the calculations, we can make comparisons between the gravity data and the atmosphere. Then, after we have enough knowledge of the data, we can start with the calculation of the atmosphere, by using only the

[1] (*Wikipedia: Luftdruck*, n.d.)
[2] (Wiese, 2011), p.1095

gravity data from the earth. While we calculate, it is exciting to know, how close the calculated atmosphere is toward the real atmosphere. In order to make this similarity apparent, we will make different tests to see the advantages and disadvantages from the different methods. At the end, we make a conclusion, how good our results are and look forward, what else could be done.

So, now we can start with the question, if we can estimate the atmosphere by using only the gravity data from a satellite.

## 1.2 Problem statement

The aim of this thesis is to solve the question, if it is possible to use the measured gravity field of the earth by the new satellite mission to estimate the atmosphere in a way, that it is usable for the atmospheric research. Also a real time estimation of the atmosphere will be tested. Because of the limit of space, we will do mostly basic tests, but further tests will be discussed also.

# Chapter 2

# Analysis of the data

## 2.1 The used data

Before the question can be solved, how to estimate the atmosphere, we have to learn more about the data we use for these calculations. In short, there are two kinds of data, on the one hand, data from the atmosphere which is later used to compare the results of the estimated atmosphere and on the other hand are the so called Daily Wise Solutions which are the simulated measured gravity field.

## 2.2 The spherical harmonic coefficients

By having a look at the Daily Wiese Solutions, there is one detail, which is worth to be discussed a bit more precisely, before we conduct further analysis. The whole data, which is used in the thesis, is given in spherical harmonic coefficients. During the analysis of the question, if it is possible to estimate the atmosphere by using the measured gravity field, it would be very useful, to know more about them. At first, the question occurs, what is behind the word spherical harmonics? The answer is easy, if we remember the potential theory, which explains that from each body, the gravitational field can be explained as a harmonic function. This harmonic function fulfils the Laplace equation and because we are on the earth which has a spherical geometry, we can write the mentioned Laplace equation by using spherical coordinates[1]:

$$r^2\Delta\Phi = r^2\frac{\partial^2\Phi}{\partial r^2} + 2r\frac{\partial\Phi}{\partial r} + \frac{\partial^2\Phi}{\partial\vartheta^2} + cot\ \vartheta\frac{\partial\Phi}{\partial\vartheta} + \frac{1}{sin^2\ \vartheta}\frac{\partial^2\Phi}{\partial\lambda^2} = 0 \tag{2.1}$$

Here is $\lambda$ the longitude, $\vartheta$ is the co-lathitude and r is the variable for the radius.

In the next step, the approach of separation:

$$\Phi(\lambda, \vartheta, r) = g(\vartheta)h(\lambda)f(r) \tag{2.2}$$

leads to the following three ordinary differential equations, each of them related to one coordinate[2]:

$$r^2\frac{d^2f}{dr^2} + 2r\frac{df}{dr} - n(n+1)f = 0 \tag{2.3}$$

[1] Chapter 2, (Sneeuw, 2015)
[2] Chapter 2 Spherical harmonics, (Sneeuw, 2015)

$$\frac{d^2h}{d\lambda^2} + m^2h = 0 \tag{2.4}$$

$$\frac{d^2g}{d\vartheta^2} + \frac{dg}{d\vartheta}cot\ \vartheta + \left(n(n+1) - \frac{m^2}{sin^2\ \vartheta}\right) g = 0 \tag{2.5}$$

In this case $m$ is the order and $n$ is the degree.

Now the differential equations can be solved, where it can be seen, that the solutions of the first two equations are the formulas[3]:

$$f(r) \in \{r^n, r^{-(n+1)}\} \tag{2.6}$$

$$h(\lambda) \in \{cos\ m\lambda, sin\ m\lambda\} \tag{2.7}$$

To solve the last differential equation, the associated Legendre functions from the first and second kind are needed, which can be seen in the following[4]:

$$g \in \{P_{n,m}(cos\ \vartheta), Q_{n,m}(cos\ \vartheta)\} \tag{2.8}$$

After solving them, it is possible to write the linear combination of $f, g, h$ in the following way[5]:

$$\Phi(\lambda, \vartheta, r) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} \alpha_{n,m} \left\{ \begin{array}{c} P_{n,m}(cos\ \vartheta) \\ Q_{n,m}(cos\ \vartheta) \end{array} \right\} \cdot \left\{ \begin{array}{c} cos\ m\lambda \\ sin\ m\lambda \end{array} \right\} \cdot \left\{ \begin{array}{c} r^n \\ r^{-(n+1)} \end{array} \right\} \tag{2.9}$$

There the arbitrary coefficients $\alpha \in R$ are solutions of the mentioned Laplace equation[6].

For the gravitational potential two properties should be fulfilled:

- for infinite distances the gravitational potential should tend to zero
- it should also be limited on the sphere

Regarding these properties, it can be noticed that $Q_{n,m}(cos\ \vartheta)$ and $r^n$ should be excluded from the model. The reason is that $Q_{n,m}(cos\ \vartheta)$ is infinite at the poles and $r^n$ is not zero in the case $r \to \infty$. Because of numerical reasons, only normalized Legendre functions will be used in the following, which can be seen below[7]:

$$\bar{P}_{n,m}(cos\ \vartheta) = P_{n,m}(cos\ \vartheta)\sqrt{(2-\delta_{m0})(2n+1)\frac{(n-m)!}{(n+m)!}} \tag{2.10}$$

[3] Chapter 2 Spherical harmonics, (Sneeuw, 2015)
[4] Chapter 2 Spherical harmonics, (Sneeuw, 2015)
[5] Chapter 2 Spherical harmonics, (Sneeuw, 2015)
[6] Chapter 2 Spherical harmonics, (Sneeuw, 2015)
[7] Chapter 2 Spherical harmonics, (Sneeuw, 2015)

Until now we have worked on the unit sphere, now, in the last step, we want to bring it to the radius $R$ and also use the mentioned normalized Legendre function. Doing that, we receive the following formula for $r \geq R$ for the gravitational potential[8]:

$$V(\lambda, \vartheta, r) = \frac{GM}{R} \sum_{n=0}^{N_{max}} \sum_{m=0}^{n} \left(\frac{R}{r}\right)^{n+1} \bar{P}_{n,m}(cos\ \vartheta)\ (\bar{C}_{n,m} cos\ m\lambda + \bar{S}_{n,m} sin\ m\lambda) \tag{2.11}$$

Because in the thesis the common letters for the degree and the order of the coefficients will be used, the formula of the Earth gravitational potential has to be rewritten $V$[9]:

$$V(r, \theta, \lambda) = \frac{GM}{R} \sum_{l=0}^{\infty} \left(\frac{R}{r}\right)^{l+1} \sum_{r=0}^{l} \bar{P}_{lm}(cos\ \theta)(\bar{C}_{lm} cos\ m\lambda + \bar{S}_{lm} sin\ m\lambda) \tag{2.12}$$

Where $l$ is the degree and $m$ is the order of the coefficients. The degree and order are in the following relation to each other:

$$l \geq m \tag{2.13}$$

As it can be seen in ( 2.12), the terms $C_{lm}$ and $S_{lm}$ which describes the amplitude of the sine and cosine terms are the mentioned coefficients. So, here is always a pair of coefficients, one for the sine term and one for the cosine term. There is also a physical interpretation of the lower spherical harmonic coefficients. One important coefficient is $\bar{C}_{0,0}$ which represents the complete mass of the earth, but because the coefficients are normalised, $\bar{C}_{0,0}$ is 1. Another important coefficient is $\bar{C}_{2,0}$, this one represents the dynamic flattening of the earth and therefore, the ellipsoid. The 7-parameter of the coordinate system is defined by the coefficients $C_{0,0}, C_{1,0}, C_{1,1}, S_{1,1}, C_{2,1}, S_{2,1}$ and $S_{2,2}$[10]. After these basics, it is to mention that the coefficients which are used here are normalised coefficients. They can expressed as in the following equation[11]

$$\left.\begin{matrix} \bar{C}_{l,m} \\ \bar{S}_{l,m} \end{matrix}\right\} = \frac{1}{2l+1} \frac{1}{MR^l} \iiint_V r^l \bar{P}_{l,m}(cos\ \theta) \left\{\begin{matrix} cos\ m\lambda \\ sin\ m\lambda \end{matrix}\right\} \rho dv \tag{2.14}$$

After this excursion in the world of SH-Coefficients, we can have a closer look in which way the gravity data is shown in the following thesis.

[8] Chapter 2: Spherical harmonics, (Sneeuw, 2015)
[9] (Sneeuw, 2006), p. 79
[10] (Sneeuw, 2006), p.92
[11] (Sneeuw, 2006), p.90

## 2.3 Representation of the potential

After learning a lot about the spherical harmonic coefficients and all basics are recapitulated, we begin now to analyse the data. But, if the data should be plotted, the way in which the density variations in the plot should be depicted has to be chosen. In this thesis, the geoid height will be used to show the variations. The geoid height is defined as the distance along the plumb line from an ellipsoid to a geoid surface. In this case, the WGS84 (World Geodetic System 1984) ellipsoid is chosen. At this point the question could occur, how the density variations can be calculated. Here the Brun's Formula come into action, which can be seen in the following ( 2.15):

$$N = \frac{T - \Delta w}{\gamma} yy \tag{2.15}$$

y

Where $N$ is the geoid undulation, $T$ is the disturbing potential, $\Delta w$ is the gravity anomaly and $\gamma$ is the normal gravity. With this it is possible to calculate the changes in the gravity, which are caused by the atmosphere. After that, all basics have been discussed, so in the next chapter it will be explained, how the so called Daily Wiese Solutions are created, which are the simulated measured gravity fields. Also, the benefits of the chosen number of satellites will be shown.

## 2.4 The benefit of two satellite pairs and the Daily Wiese Solutions

### 2.4.1 The satellite constellation

Before the analysis of the question can start, the advantages of the proposed satellite constellation will be analysed. As Mr. Wiese mentioned in his paper (Wiese, 2011), the concept of using one pair of satellites in a polar orbit, while another pair flies in an orbit with a lower inclination, has several advantages. This concept can provide information in the East-West direction, which are missing in the data, provided by the current missions. Another point is that the effect of the stripes which appear longitudinal in the current solutions can be presumably reduced(Bender, 2008). It is also to mention that these satellites, in opposite to the current mission, could fly drag free, which means that the error can be reduced, which occurs during the measurement of forces which are non-conservative(Loomis, 2009). Drag-free satellites have the property that their path is geodesicy, which mean they use the shortest path between two points, therefore they are only affected by gravity. Thus, forces from non-gravitational effects have no influence on them, for example the drag which occurs from the remaining atmosphere in the lower altitude of the geodesic satellites. Also the effects of the sun like solar wind are reduced or eliminated [12]. Besides the advantage of the drag free flight, the new technology also gives an improvement in the area of measurement instruments. Instead of the microwave ranging instrument, which was used in the GRACE satellites[13], we can now use a laser interferometer, which provides us with a high improvement in accuracy. Therefore, the error of the

[12](*Wikipedia: Zero drag satellite*, n.d.)
[13](*Wikipedia: Gravity Recovery and Climate Experiment*, n.d.)

measurement is lower, compared to the error of, for example, the temporal aliasing, by measuring the hydrology and also to the errors in the models for the atmosphere or oceans[14], which clearly can be seen in the following figure, which is published in (Wiese, 2011):

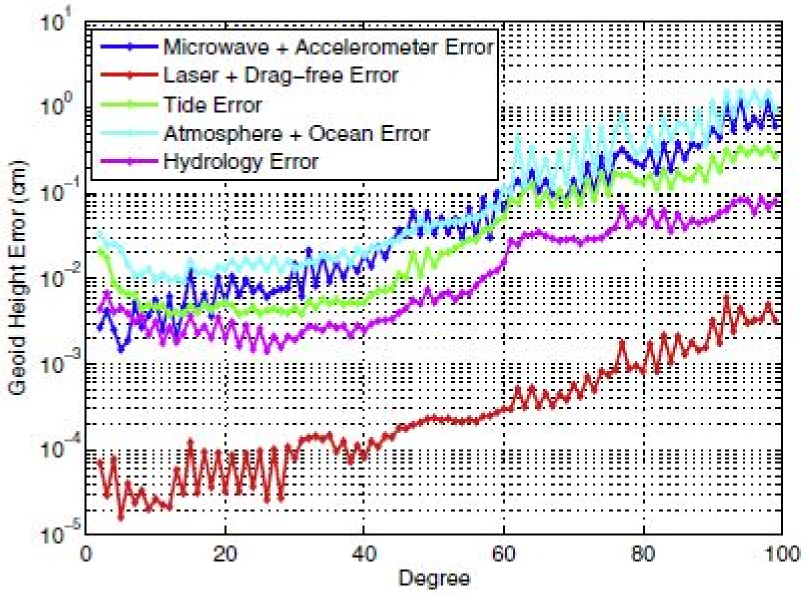



***Figure 2.1:*** *Benefit of the Laser interferometer, compared to microwave ranging instrument and error sources*

As it can be seen in figure 2.1, the error of the laser interferometer is clearly smaller than the other errors. Thus, the main error sources are the hydrology and atmosphere. These two sources of error have the *benefit* that they can be easily reduced, while a measurement error is quite harder to reduce. It also is helpful to know for the further analysis that the satellites which are simulated in the used data have a repeat period of 11 days, hence, the daily solutions are organised in 11 day groups, which will be presented later. Now, all basics about the satellites have been discussed, which were simulated to estimate the daily solutions. It could be seen that by using them it is already possible to reduce some error sources, but there are still errors left. The main errors are caused by the atmosphere, which we want to analyse, and the ocean.

### 2.4.2 Main errors and how to eliminate them

At this point, the Daily Wiese Solutions come into play. The main aim of the daily solution is, to estimate gravity fields with a low expansion degree for a short time, in this case one day, to reduce the effects which are based on the 'temporal aliasing errors' (Wiese, 2011). To understand better what lies behind the temporal aliasing errors, it is helpful to take a look at the Nyquist theorem. The error occurs, if the sample frequency is smaller than two times the highest frequency of signals contingent. The aliasing effect is created by frequencies, which occur by under sampling, which means the previous Nyquist theorem is not maintained, or it can

[14](Wiese, 2011)

occur if the wanted signal is covered by a noise signal. These frequencies create now interfering signals in the wanted frequency domain [15]. According to Wiese (Wiese, 2011), the signals which have to be analysed can be separated in three categories, the one with no presumptive information, signals like hydrology, which have a full alias and thirdly, signals, where a model is available, with which the effect of the aliasing can be reduced, for example from the atmosphere or the tides (Wiese, 2011). This already shows the problem that good models are required to reduce the error of the temporal aliasing, which is actually the main error source and occurs in the models of the signals from the atmosphere and the ocean, 'also referred to as error in the atmosphere and ocean dealiasing (AOD) product' (Wiese, 2011). It is also helpful to know that most of the AOD error occur at low degrees, which leads to the idea that it is possible to reduce the effect on the gravity fields of the temporal aliasing errors, by estimating gravity fields which have also a low degree at a short period of time. This can be used to estimate errors by using atmospheric and ocean models,[16], which is done by estimating the Daily Wiese Solutions. To understand that better it is helpful to know, that past degree 45 the models of the atmosphere and the ocean are close to 100% error[17].

### 2.4.3 Estimating the gravity fields

Now, after the benefits from the usage of the daily solutions and the special satellite concept are shown, the question occurs, how can they be simulated? Before a mission can start, it is very important to have an idea, how the gravity fields, which would be measured by the satellites, will look like. Doing that, the NASA Goddard Space Flight Center provide the software program GEODYN which can determine the orbit of satellites[18]. As it is explained by Mr. Wiese(Wiese, 2011), the using of this software creates two cases. The so called true case regarding the errors in the measurement system and the satellite position, thus it is comparable to a real satellite mission. Also in this case there is one model of the forces included, which affects the satellite. Beside this case, there is also the nominal case, which includes several different models of the forces, and represents the best estimation of the simulated satellite mission. The signals of the hydrology as well as the variations of the ice mass can be recovered by the simulation. By using the set of force models from the nominal case, tidal, atmospheric and oceanic effects are mitigated from the solution. Here, it is also to mention that in the static gravity field, errors will not be considered. But with all these cases and models, how can the tidal, atmospheric and oceanic error occur? These errors are defined by calculating the differences between the two sets of models(Wiese, 2011). How the estimation of the satellite orbit works, is described here (Wiese, 2011):

> The satellite orbits are integrated through the truth and nominal force models, and a time series of range-rate residuals is created. The estimation is done in three steps: first, the positions and velocities of the spacecraft are estimated using the range-rate residuals along with residuals in the satellite position. The satellite orbits are then integrated again through the nominal force models using the updated spacecraft state, and a new set

[15](*Wikipedia: Alias-Effekt*, n.d.)
[16]e.g. (Wiese, 2011)
[17]e.g. (Wiese, 2011)
[18](Palvis, 2010)

of range-rate residuals is formed. The final step estimates the spherical harmonic coefficients of the gravity field, along with corrections to the state of the spacecraft using only range-rate data by parametrizing the state of the spacecraft in baseline parameters and constraining nine of the twelve state parameters, as described in(Rowlands, 2002).

Now the question arises, how the estimation process works? It can be said that so called 'daily arcs' are used(Wiese, 2011), in which all observations are agglomerated over one day, and the result of the agglomeration can be used in the normal equation. In this case, an 11 day assessment of the gravity field is needed, so that SOLVE can be used, which is a solver of linear systems, to combine 11 one day arcs(Ullman, 1997)[19]. The mentioned program SOLVE is also contributed from the Goddard Space Flight Center and is a complementation of GEODYN. If we have now a closer look at SOLVE, we can see that it specifies between two kinds of parameters. The first ones are the so called arc parameters, which are parameters about the elements which change during the simulation, for example the position of the spacecraft. The other parameters are global parameters, which describe, in the opposite to the former ones, factors which do not change during the simulation, like the gravity field, which is described by spherical harmonic coefficients. After a modification, it is now possible, to determine spherical harmonic coefficients, which are used in SOLVE, as arc parameters. Thus we can now evaluate every arc parameter along the position of the satellites[20]. After the theoretical explanation, we will now have a look at the calculations, which have been described in an abstract way so far. In the first step, the accumulation of normal equations over every arc will be done, which are given by (Tapley, 2004b)[21]:

$$Ax = b \tag{2.16}$$

Where $A$ and $b$ are:

$$A = H^T W H + \bar{P}_0^{-1} \tag{2.17}$$

$$b = H^T W y + \bar{P}_0^{-1} \bar{x}_0 \tag{2.18}$$

Where:

- $x$ = vector with unknown elements
- $H$ = Matrix which contains the partial derivatives of our observations regarding to the unknown elements
- $W$ = the weight matrix regarding the observations
- $\bar{P}_0$ = A priori matrix of the covariances
- $\bar{x}_0$ = A priori information for the x vector

According to David Wiese, $W$ defined in a way that all used measurements for the equation have an equal weight and the used standard deviation is equal to the standard deviation of the noise from the measurement(Wiese, 2011). Also $\bar{P}_0$ provides no a priori information, which has the advantage that there is a big flexibility in this solution. At least, it should be mentioned that $\bar{x}_0$ is received from the status parameters which are defined in the mentioned nominal case.

[19](Wiese, 2011)
[20](Wiese, 2011)
[21](Wiese, 2011)

By using ( 2.16), it is possible to apply a block inversion (Ullman, 1997). During the following estimation, the mentioned arc and the global parameters are separated. Doing this, we have the following equation:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{2.19}$$

As it can be seen, there are two elements in the $x$ and $b$ vectors, the upper one represents the arc parameters while the lower one represents the global parameters. Hence, the goal is reached to separate the two parameters. Now the $x$ matrix should be reduced from Eq. ( 2.19), which happens in two steps. First $A_{11}^{-1}$ will be multiplied with the first row and in a second step the first row will be multiplied with $A_{21}A_{11}^{-1}$ and the result is subtracted from the second row. Doing this, the following equation will be reached[22]:

$$\begin{pmatrix} A_{11}^{-1}A_{11} & A_{11}^{-1}A_{12} \\ A_{21} - A_{11}A_{21}A_{11}^{-1} & A_{22} - A_{12}A_{21}A_{11}^{-1} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A_{11}^{-1}b_1 \\ b_2 - A_{21}A_{11}^{-1}b_1 \end{pmatrix} \tag{2.20}$$

This equation can now be reduced to the following one:

$$\begin{pmatrix} I & A_{11}^{-1}A_{12} \\ 0 & \bar{A}_{22} \end{pmatrix} \times \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} A_{11}^{-1}b_1 \\ b_2 - A_{21}A_{11}^{-1}b_1 \end{pmatrix} \tag{2.21}$$

Where $\bar{A}_{22} = A_{22} - A_{12}A_{21}A_{11}^{-1}$. In a last step, the equation can be solved to get the arc parameters, which are the parameters from the first column[23], as already mentioned:

$$x_1 = A_{11}^{-1}\left[\left(I + A_{12}\bar{A}_{22}^{-1}A_{21}A_{11}^{-1}\right)b_1 - A_{12}\bar{A}_{22}^{-1}b_2\right] \tag{2.22}$$

At this point in ( 2.22), it can be defined, which degree and order the arc parameters of the gravity field should have. After that, the equation can be solved for every arc and as a result eleven one day estimated gravity fields will be received with a low degree and one eleven-day-estimation of the fields in a higher degree. By using the low degree and order gravity fields, mass variations can be estimated which have a high frequency and by doing that the temporary aliasing errors can be reduced. After that it is possible to improve the higher degrees. With this step, the whole process of the estimation of the daily solutions is explained, but as it can be seen later, there are also so called eleven day mean fields, which have a higher resolution than the daily solutions. To create these fields, SOLVE can be used again and only the needed degree of the wanted gravity field has to be changed[24].

[22](Wiese, 2011)
[23](Wiese, 2011)
[24](Wiese, 2011)

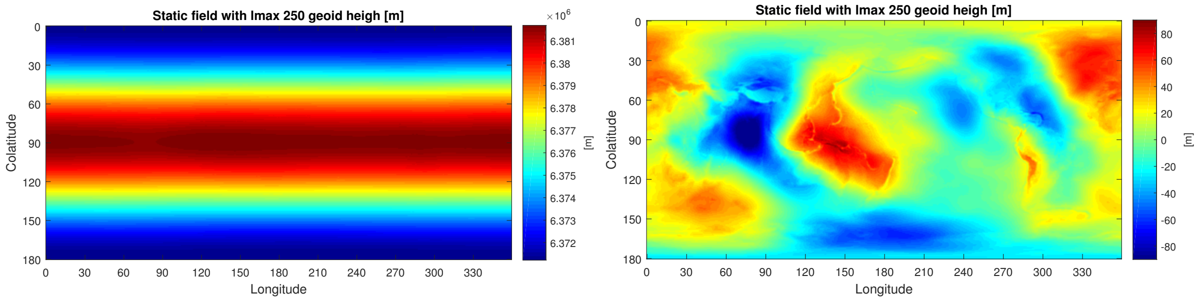
## 2.5 The daily solutions

### 2.5.1 A closer look at the given data

After the previous chapter, it is now clear, how the daily solutions are calculated. So, now the analysis of the data can start. It is to mention that at the beginning only the gravity field will be discussed, which will be shown in different variations and there are already some possibilities to emphasise short term variations. In the next step, further comparisons can be done. This is an important step to get a first idea how the data looks and which possibilities present themselves for the analysis in the next chapters. At first, the Daily Wiese Solutions will be shown, which are the calculated solutions of the gravity field for each day. The maximum degree of the SH-coefficients is twenty, which means they have a lower spatial resolution compared to the atmospheric data, which can be seen later. The data are available for 363 of the year 1996, which are organised in eleven day periods. The organisation of the days is also the reason, why only 363 days are available. It to mention, that they are tidal free, which means that during the calculation, the tidal effects were removed. To have a first impression, how the Daily Wiese Solutions look like, they will be plotted now, but by doing this, the first problem occurs. The daily solutions contain a so called static field which has a larger magnitude than the daily solutions. Therefore, to put it differently, the wanted informations are covered by the static field. To solve the problem, it is possible to subtract the static field from the daily solutions to make the temporary information visible.

After this is done, the data can be plotted, by using the functions from the SH-Bundle from the Institute of Geodesy[25]. The field, which is created by one of the functions from the Bundle, is plotted on a grid, which is determined from the field of coefficients. Now, the knowledge of the coefficients is needed, which informs us that the lower coefficients describe the shape of the earth. This is a vital information, because we have to decide, if the WGS84 Ellipsoid from the given field should be subtracted or not. For all data which is in the original configuration, it is very helpful to use this, to eliminate effects which cover the wanted details. The overlay is caused by the big variations like the flattening of the earth, which are described by the lower coefficients cover the wanted details, which are represented in smaller variations. This effect is shown in the following on the left side for the static field with a maximum degree of 250, if the WGS84 Ellipsoid is subtracted, the gravity field from the static field can be seen clearly. Which is also shown below on the right side of the figure below:

[25](*SHBundle: http://www.gis.uni-stuttgart.de/research/projects/bundles/*, n.d.)

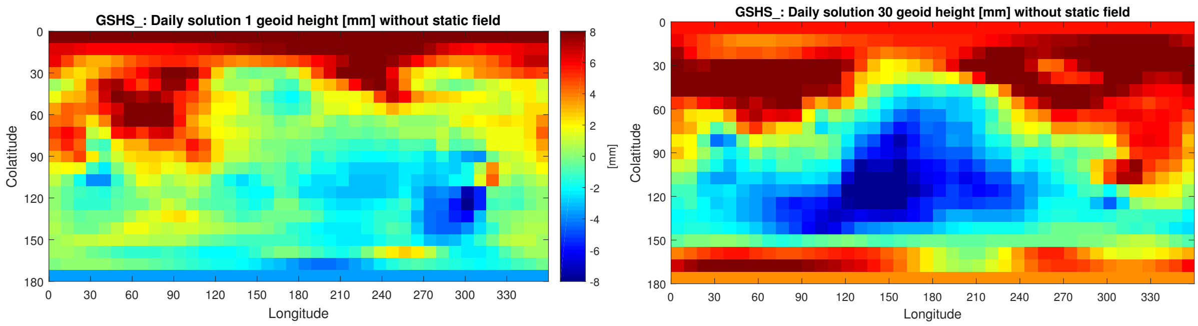



(a) The static field without the subtraction of the WGS 84 Ellipsoid

(b) The static field with the subtraction of the WGS 84 Ellipsoid

***Figure 2.2:*** *Comparison between the output, if the WGS 84 Ellipsoid is subtracted from the static field or not*

The higher coefficients represent the smaller variations, but, if the ellipsoid is not eliminated, the small variations are covered by the ellipsoid. The high values in the middle of the left image of figure 2.2 show the semi major axis of the ellipsoid, the lower values at the poles are the semi minor axis. If we want to eliminate the WGS84 Ellipsoid, the lower coefficients were subtracted to eliminate the ellipsoid and to show the smaller variations. In the following the daily solutions of the first and 30th day can be seen without the static field:

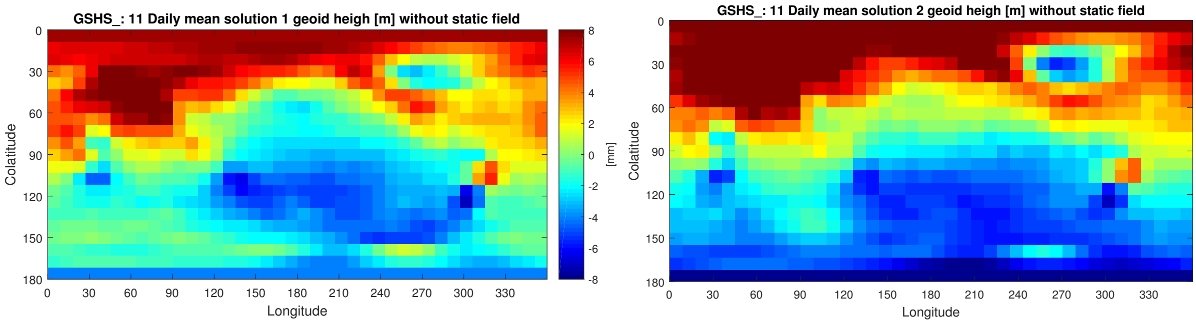



(a) Daily solution without static field for the first day

(b) Daily solution without static field for the 30th day

***Figure 2.3:*** *Example of two daily solution fields without the static field*

Here, the variations of the gravity field can be seen, which occur during the measured time. Now, after the static field is subtracted from the daily solution to eliminate the dominating effects, the effects of long term variations also can be reduced by subtracting the mean field from the daily solutions. It can be assumed that the atmosphere changes very fast and in this way, disruptive effects from the long term variations which maybe cover the effects of the atmosphere, are reduced. This field is given until degree 120, so it has also a finer spatial resolution than the daily solution. In the following figure 2.4, the reduced first two mean fields can be seen:

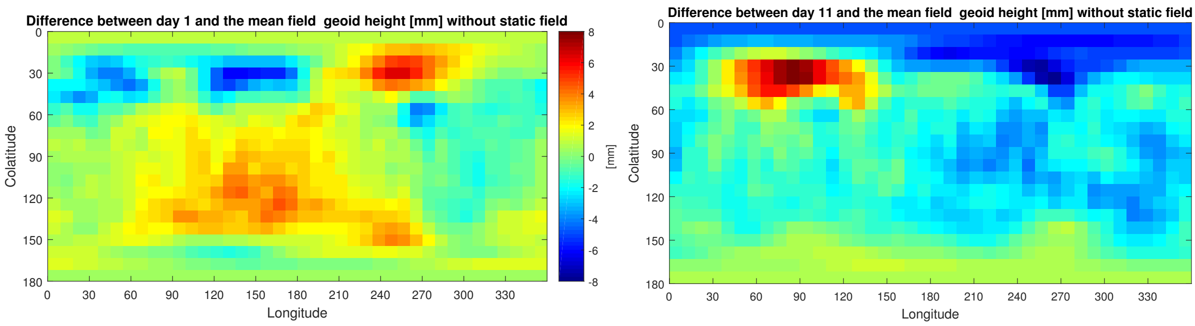



(a) The mean field of the first day to the 11th day until degree 20 without the static field

(b) The mean field of the 12th day to the 22th day until degree 20 without the static field

***Figure 2.4:*** *Example of two mean fields without the static field*

A look at the figure 2.4 shows that there are some similarities between the two fields, which would indicate that some effects appear in both timespans which are covered by the two mean fields. However, in the right mean field the positive pattern appears to be much stronger compared to the left one. Also, the negative pattern increased. Now, the short time variations can be seen, after the mean field is subtracted from the daily solutions:

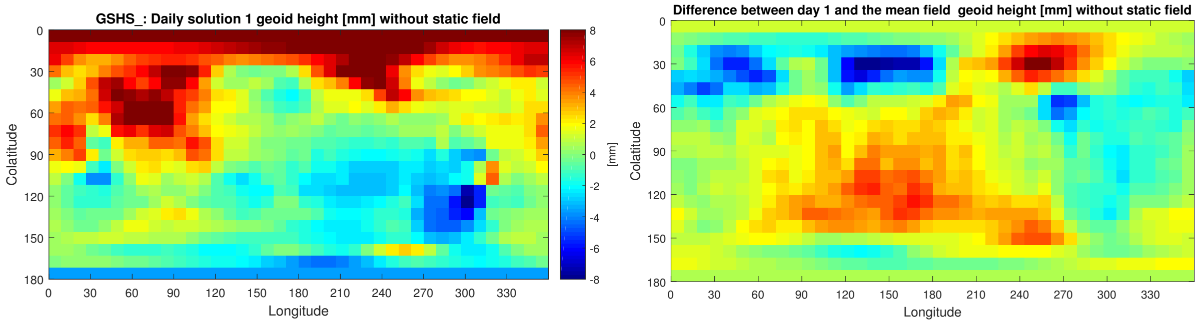



(a) The daily solution of the first day subtracted by the mean field

(b) The daily solution of the 30th day subtracted by the mean field

***Figure 2.5:*** *Examples of two daily solutions without the mean field*

This method offers a big change compared to the previous comparison in figure 2.4. Here in figure 2.5 are now smaller structures visible and also pattern appear which could not be seen in the previous comparison. To see this changes better in the following the first day is shown with and without the eleven day mean field:




(a) Daily solution of the first day (b) Daily solution reduced with the 11 day mean

***Figure 2.6:*** *Comparison between the daily solutions with and without the subtraction of the 11 day mean field*

The comparison in figure 2.6 shows some big variations between these graphics, which lead to the awareness that there are actually many short time variations. For further comparisons with the atmosphere, it is interesting to compare both fields with the atmosphere, to see the similarities. But at the moment, it is difficult to imagine, where on the earth which signals occur, for that, a map of the world will be now overlapped by the daily solutions, reduced by the eleven day mean field:

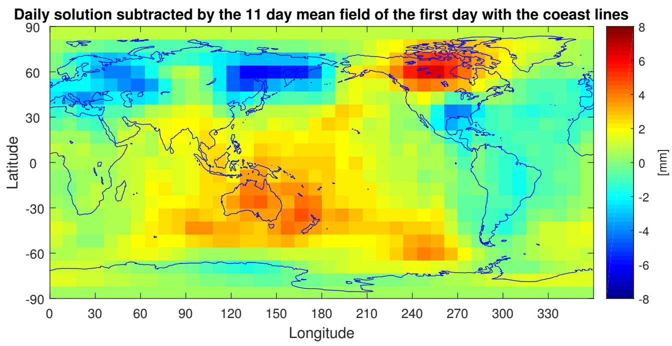
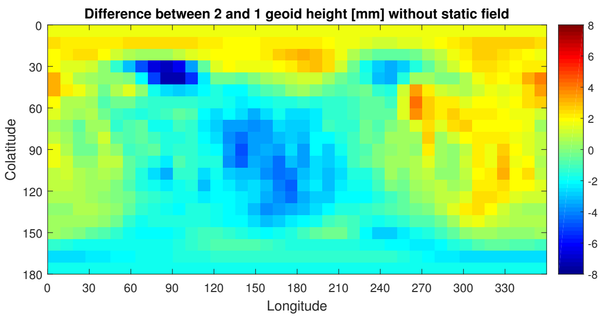
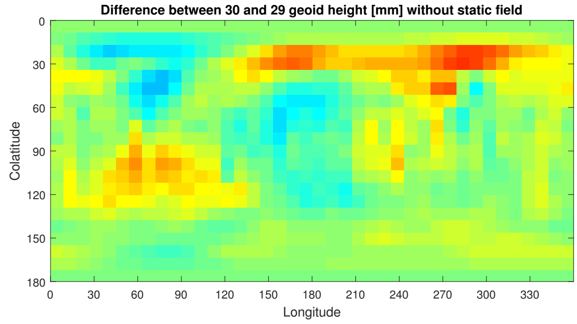



***Figure 2.7:*** *Daily solution subtracted by the 11 day mean field, covered by a map*

As it can be seen, there are strong negative pattern over Asia and strong positive pattern over Canada. Europe is covered by a weak negative pattern as well as the south USA, while Australia has a medium strong positive pattern. It can be imagined, at which part of the earth the different pattern occur. In the next step, we try another possibility to show the variations by subtracting the gravity fields of two following days. Hence, only the differences, the variations, which occur between the two days, can be seen. This is done in the figure below with the first and second day and also with the 29th and the 30th day. We can also mention that the subtraction of two fields happens by subtracting two fields in the spherical harmonic domain.




(a) Difference between the daily solution of the first and second day




(b) Difference between the daily solution of the 30th and the 29th day

***Figure 2.8:*** *Differences between two subsequent daily solutions*

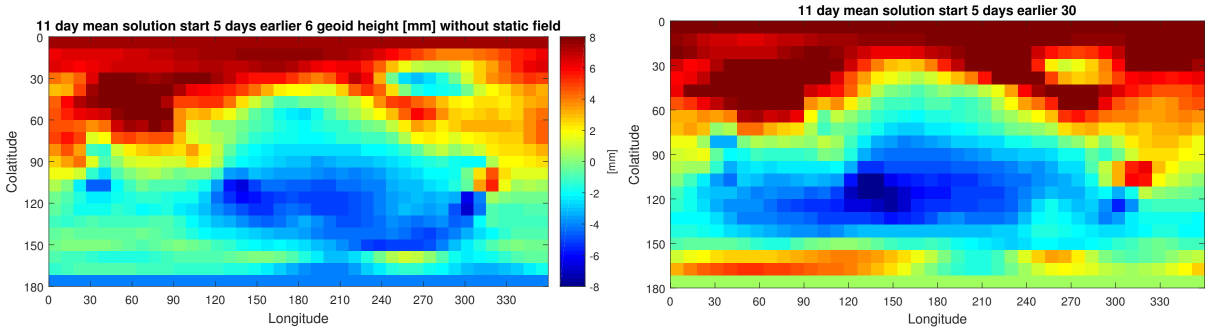
It becomes apparent that there are also big variations between two days, which mean that there are also many changes in the daily solutions. One possibility is that some of these changes are created by the effects of the atmosphere. However, regarding the main goal of the these, the question occurs at this point,if all effects are caused by the atmosphere or if there also additional effects from the tides, which are not completely eliminated.

Now after we had a look at the given data and the using of the given 11 day mean fields, we should regard one point. By subtracting the 11 day mean field, we subtract for 11 days the same field, for the next 11 days another mean field, thus we create an 11 day periodicity in our data, which could at the latest create wrong signals if we have a look at the frequency domain. So it is necessary to avoid this effects, but as we have seen, the subtraction of the fields could have some improvements in eliminating long term signals. So now we should search an alternative, which provides the same or even better results without creating the 11 day frequency.

### 2.5.2 Improving the daily solutions

One way to improve the daily solutions without using the given daily solutions is, that we use a moving average filter. Though we use also a 11 day mean field, but this time the field will be calculated for each day. By doing this, we start 5 days earlier and use also the five days after the day we want to improve. By doing this, we avoid the 11 day periodicity because it is calculated for each single day. Here and for all follow calculation it is to know that because of the using of the moving average, we only can use the days 6 to 358, because of the fact that 5 days before and 5 days after the actual day are used, thus we have 353 days for the calculations.

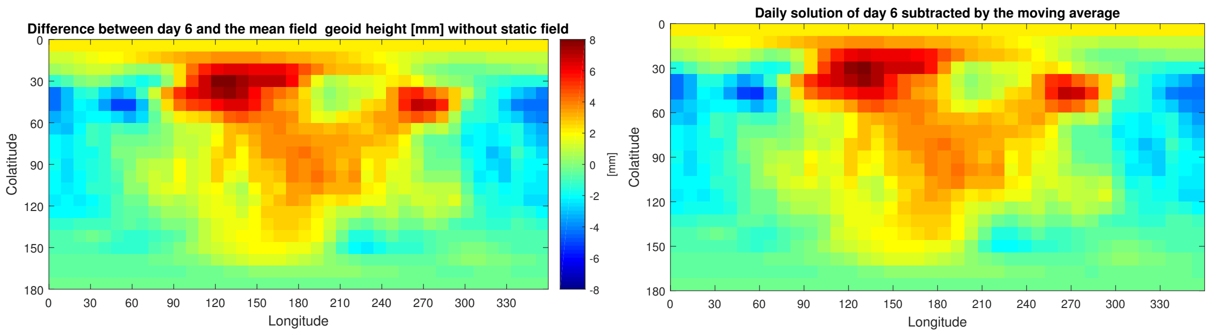
At first we will have a look at the moving average of day 6 and 30:




(a) Calculated moving average field for the 6th day (b) Calculated moving average field for the 30th day

*Figure 2.9: Example of the calculated moving average fields for the 6th and 30th day*

The fields in figure 2.5.2 are created by calculating the mean of 11 following fields, so short term values are reduced and only the long term variations can be seen. This explains why we can see here several similarities between the moving average field from the day 6 and 30. Now we can subtract the daily solutions from the calculated moving average field to see the improvement. Doing this we can compare the daily solution which is subtracted from the given 11 day mean field and the daily solution which is subtracted by the moving average:




(a) Daily solution of the 6th day subtracted by the 11 day mean field (b) Daily solution of the 6th day subtracted by the moving average

*Figure 2.10: Comparison of the daily solution of the 6th day subtracted by the 11 day mean field and the moving average*

Because the given 11 day mean field should be identical with the calculated moving average, we can see in 2.5.2 that the daily solution subtracted by the given 11 day mean field shows identical pattern with the daily solution subtracted with the calculated moving average field. With this it is possible to verify the result, so we can no proceed with the analysis.

For the following work we will now only use the daily solutions with the subtracted moving average. In the next step, we first have a look at the atmospheric data, which are also given. But before we do that, we will have a closer look on the signal pattern from the daily solutions for a whole year and also for the seasons. This will be done by calculating the Root Mean Square (RMS) for each pixel with the following formula:

$$RMS = \sqrt{\frac{\sum_i N_{i,k}^2}{\bar{N}}} \qquad \bar{N} \text{ - number of elements} \tag{2.23}$$

In a first step we will have a look at the comparison of the RMS map from the daily solutions, which are subtracted by the static field as well as the moving average, for the whole year:

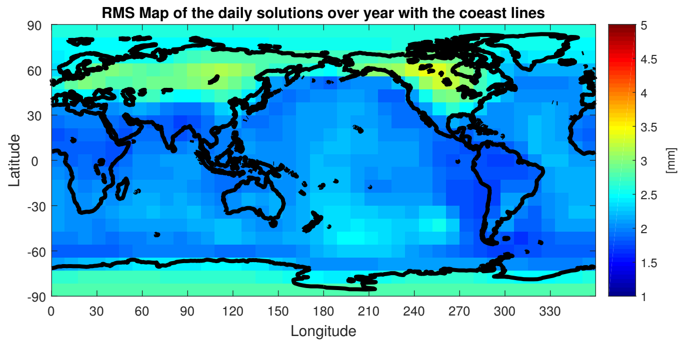
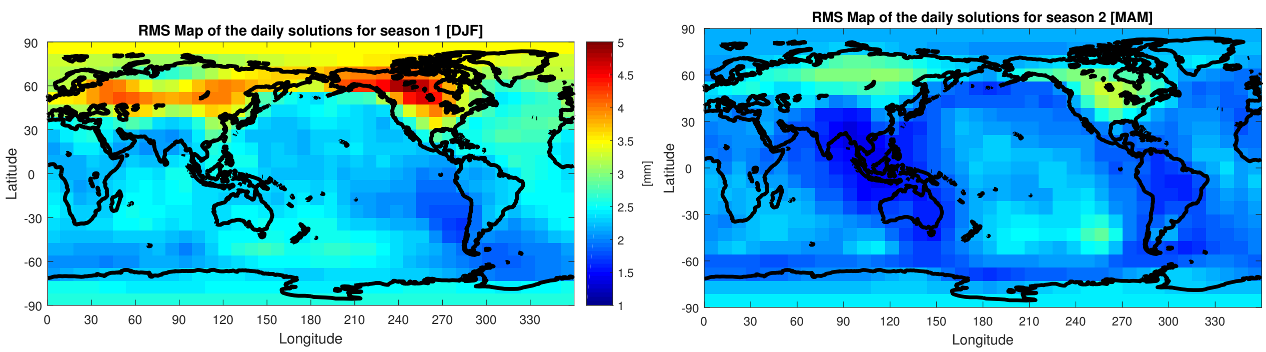



***Figure 2.11:*** *RMS map from the daily solutions over the whole year*

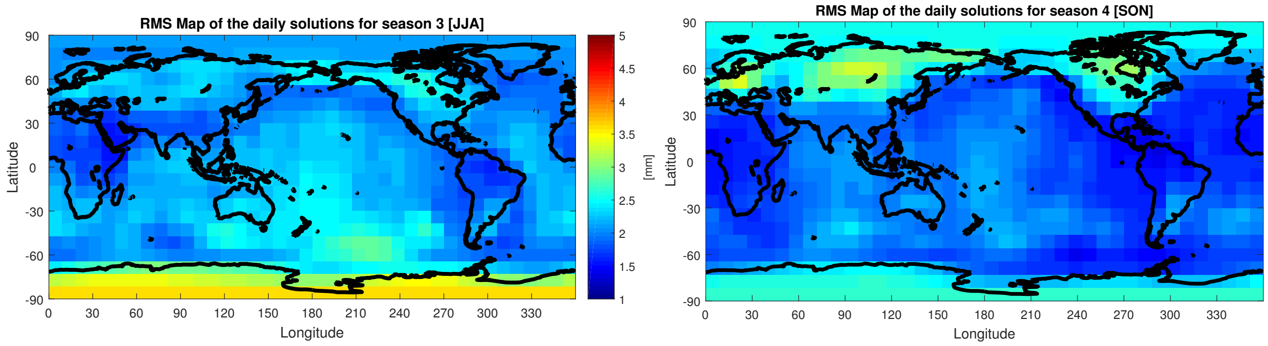
As it can be seen in figure 2.11, there are positive patterns above Russia and North America, as well as Antarctica. There is also a weak negative area above the Arctic but between the mentioned positive fields are negative fields. This way makes it easy to see that the main energy is in both fields above Asia and North America in the northern hemisphere and also above the Antarctica. To have a better impression how the signal changes, we will have now a look at the atmospheric seasons which are December - January - February as season one and March - April - May as season two which can both seen in figure 2.12 and June - July - August as season three and September - October and November as season four, which are both shown in fig. 2.13:




(a) RMS map from the daily solutions for season one

(b) RMS map from the daily solutions for season two

***Figure 2.12:*** *RMS maps of the seasons one and two*




(a) RMS map from the daily solutions for season three

(b) RMS map from the daily solutions for season four

***Figure 2.13:*** *RMS maps of the seasons three and four*

In figure 2.12 it can be seen in the RMS image of the first season that there are strong positive signals over North America and less stronger positive signals over Russia and China and the North Pole. In the second season, positive signals still can be seen over North America and Russia, but now weaker. In the third season, there is a strong positive signal around the South Pole.This explains why there are strong signals in the south in figure 2.11. In the northern hemisphere only negative pattern can be seen. At least in season four, the transition to the first season can be seen, because now we can see that the signals in the south are more weak, and the signals in Russia and North America become stronger. As a conclusion, it can be said that the strongest signals happen in the winter (season one) and the summer (season three). During these seasons, the strong signals are clear separated to the north in the winter and to the south in the summer, in the two other seasons, the signals are weaker and occur in the north and the south. But if we have a look at the meteorological seasons of the southern hemisphere, we can see that the winter is, per definition, from the 21th June until the 22th or 23th September[26], which matches with the third season. If we imagine the weather in the winter, where we have a lot of snow or rain, it can be assumed that there is a big amount of mass on the earth, which can be measured by the satellite. Therefore, in the winter in the northern hemisphere, strong signals appear in the north, while in the winter in the southern hemisphere the strong signals are in the south. In spring and autumn, we can see changes in the areas, so it is already possible to imagine the effects of the seasons only by using the daily solution. Keeping the aim of this thesis in mind, to estimate the atmosphere by using daily solutions, we now have made a first step towards the estimation. Now, we have a good idea, how the daily solutions are created and how they look like. It was also possible to assume the first similarities between the effects of the atmosphere and the daily solutions. Now we will have a look at the atmosphere.

## 2.6 The atmosphere

To analyse the influence of the atmosphere on the gravity field, we will use an atmospheric gravity model which is given also for a whole year, and has a maximum degree of 180. The speciality of this data set is that the fields are given in time steps of 6 hours. The atmospheric gravity model which is used in this work is part of the ESA's Earth System Model. Here it is

[26](*Wikipedia: Südhalbkugel*, n.d.)

to mention, that it has an inverse barometrically corrected atmospheric mass variability. The summarized effect of the hydrostatic pressure of the atmosphere and the oceanic masses which are situated along a plumb line above a position of the ocean sea floor is commonly regarded as the ocean bottom pressure. Here the problem occur, that the atmospheric and oceanic contributions to the described effect of ocean bottom pressure are high correlated and by summarizing them they largely cancel each other. The correlation is removed here by applying the so called IB-correction (Inverse barometric correction), too the oceanic components. With this it is possible to see the atmospheric and oceanic components as independent contributions to the time variable gravity field(Dobslaw, 2014). At first, we can plot the atmospheric gravity field for the whole first day without any further modification:

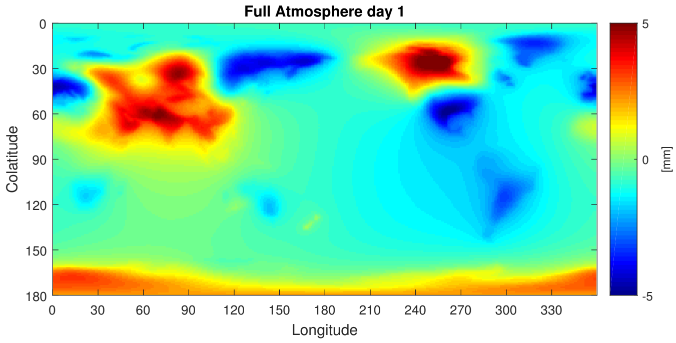



***Figure 2.14:*** *The atmospheric gravity field during the first day until degree 180*

As it can be seen in fig. 2.14, some familiar patterns are already possible, but because of the higher degree of the field, we also can see much more smaller structures. To simplify the comparison, in a first step the degree of the gravity field will be changed until it matches the degree of the daily solutions. This can be done as previous described, by using the $|c \setminus s|$ format and reducing the degree into the needed size. The result, also for day one, can be seen below:

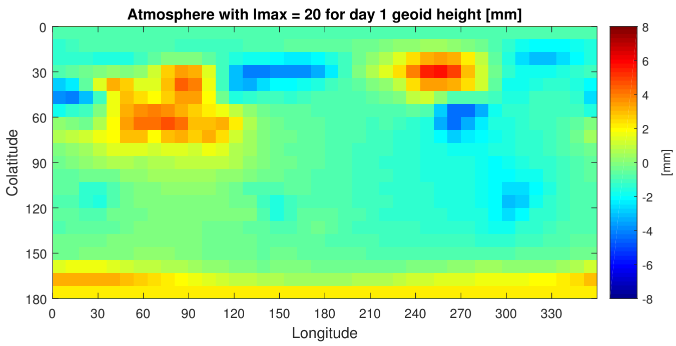
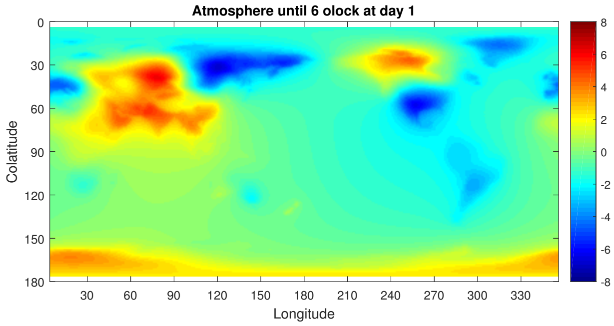
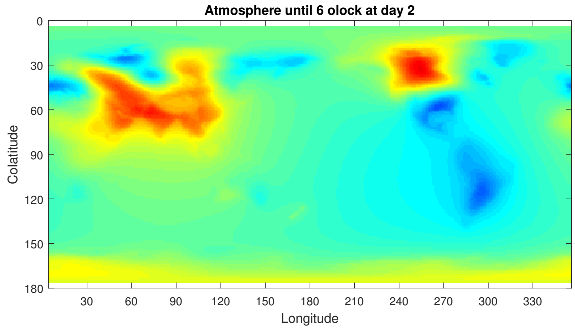



***Figure 2.15:*** *The atmospheric gravity field during the first day until degree 20*

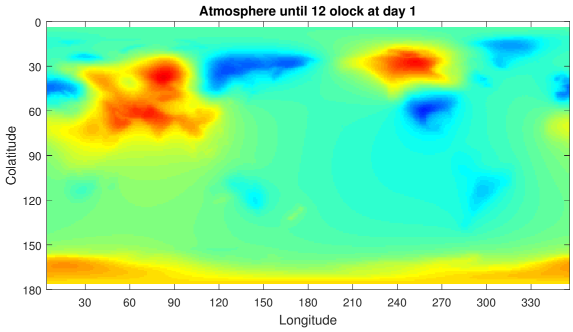
As expected, the small structures in fig. 2.15 are now eliminated, but it is now possible to compare it with the daily solutions. Now, we only saw the summary of the single 6 hour fields. To have an idea, how the single fields look like, we can plot them for two days:



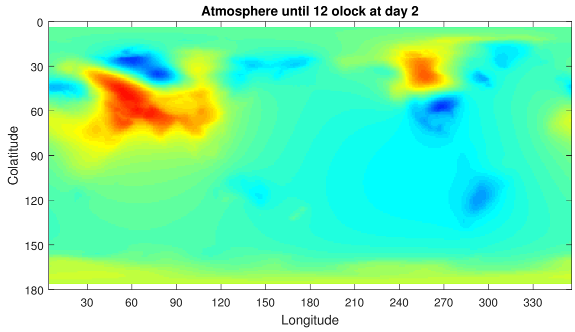

(a) Atmosphere from 0 - 6 o'clock at day one



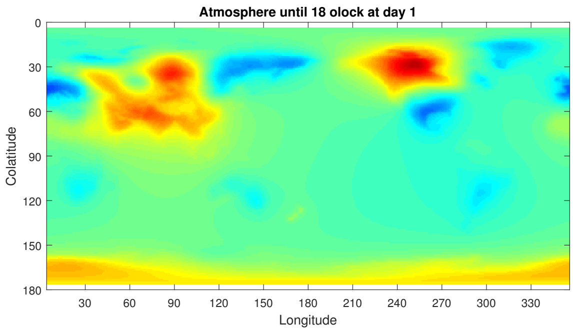

(b) Atmosphere from 0 - 6 o'clock at day two



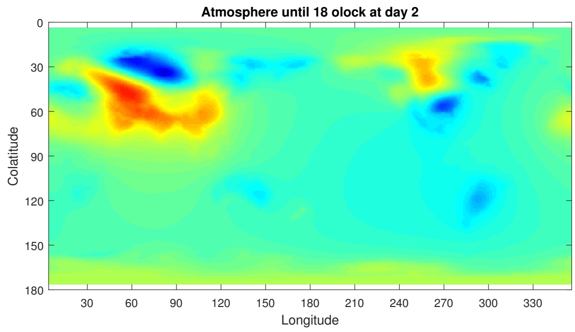

(c) Atmosphere from 6 - 12 o'clock at day one



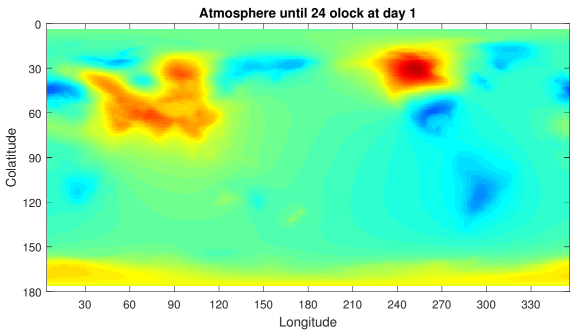

(d) Atmosphere from 6 - 12 o'clock at day two



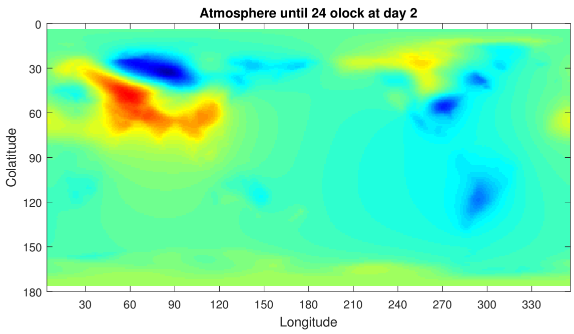

(e) Atmosphere from 12 - 18 o'clock at day one




(f) Atmosphere from 12 - 18 o'clock at day two



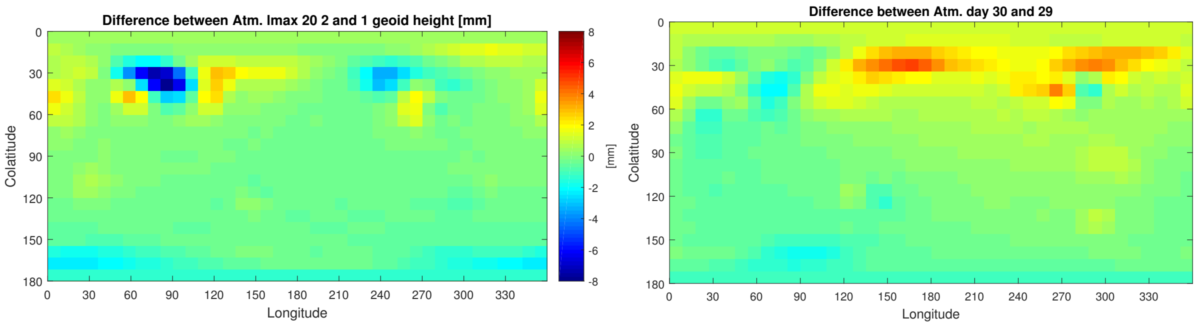

(g) Atmosphere from 18 - 24 o'clock at day one




(h) Atmosphere from 18 - 24 o'clock at day two

***Figure 2.16:*** *Epochs of the atmospheric gravity field for the first two days*

In figure 2.16, it becomes obvious, that there are already differences in the atmospheric gravity field during the day. For example, the positive pattern on the upper right side becomes stronger and bigger until the field of 6 o'clock of the second day, then it becomes weaker. On the upper left side a positive field occurs during these days and becomes stronger. Of course, it is also possible to make the changes visible by calculating the differences of two consecutive days as we have done it before with the daily solutions. The idea behind this calculation is that by calculating the differences, we can distinguish the short time variations, because the subtraction should eliminate the long term variations. In the following, the results of the differences between the second and the first day can be seen, and also between the 30th and 29th day.

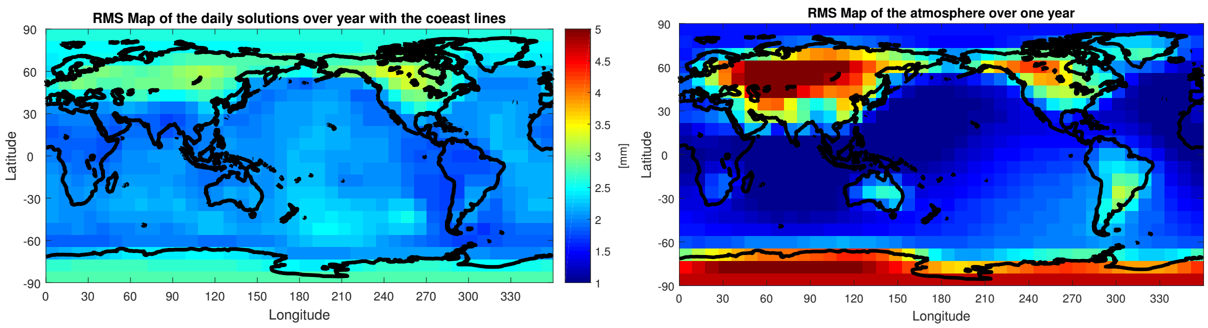



(a) Changes in the atmospheric gravity field between the first and the second day

(b) Changes in the atmospheric gravity field between the 29th and the 30th day

***Figure 2.17:*** *Comparison of the changes between two subsequent atmospheric gravity fields*

It shows that there are also big variations, which probably could be seen in the measured gravity fields. We can conclude that both, the atmospheric gravity fields and the daily solutions show big variations in short time intervals, which were made obvious by calculating the differences. We can use this fact in the next chapter to estimate the atmospheric gravity fields. At this point, we can start with the comparisons between the daily solutions and the gravity fields of the atmosphere, by comparing their RMS maps. We have already seen the RMS maps of the daily solutions and have an impression of how they look. Now, we can see the comparison to the atmospheric gravity field. In a first step, we will have a look at the comparison of the RMS maps from the daily solutions and the atmospheric gravity fields for a whole year:

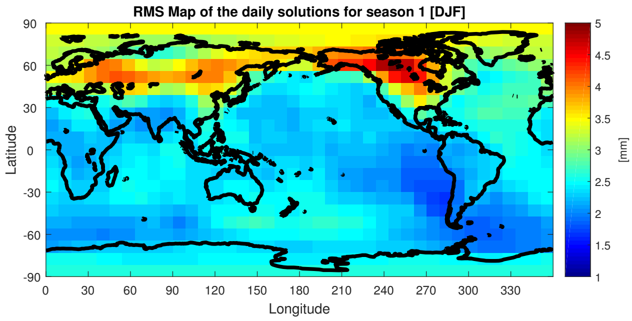



(a) RMS map from the daily solutions over the whole year

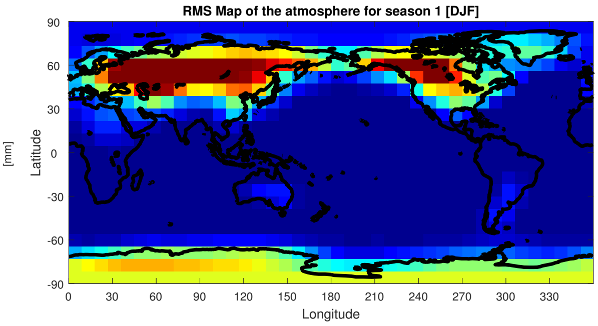
(b) RMS map from the atmospheric gravity fields over the whole year

***Figure 2.18:*** *Comparison of the RMS maps for a whole year*

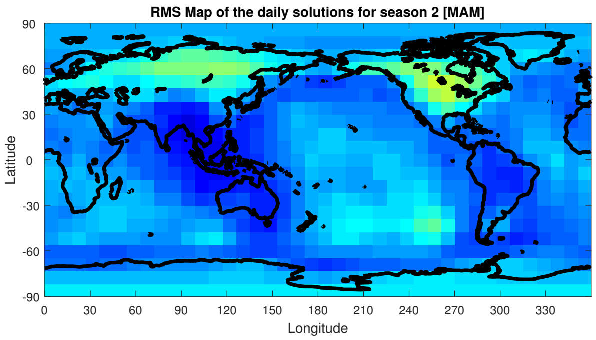
In the figure 2.18 it can be seen that there are already some similarities, but additionally, we have to mention that the signals from the atmospheric gravity fields are much stronger than the signals from the daily solutions. However, the areas, where they both have strong positive patterns match with each other. As a first assumption, it can be said that the measured gravity field shows only the effects of the atmospheric gravity fields, which means that it represents the strength of the gravity fields of the atmosphere in a moderate way. But, it is a good sign that we have here a match between the patterns. Therefore, it is possible to compare the RMS maps of the seasons from the two datasets, which we already saw from the daily solutions. Now, we can compare them with the RMS maps from the atmospheric gravity fields.



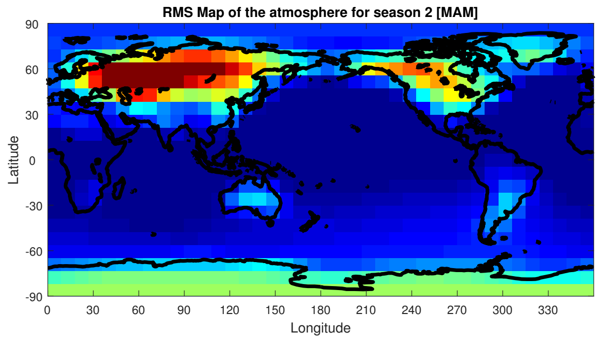

(a) RMS map from the daily solutions for season one



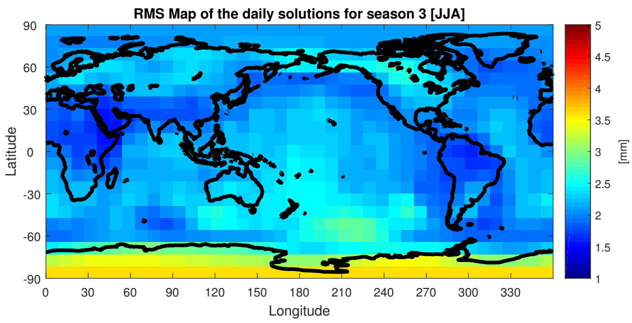

(b) RMS map from the atmospheric gravity fields for season one



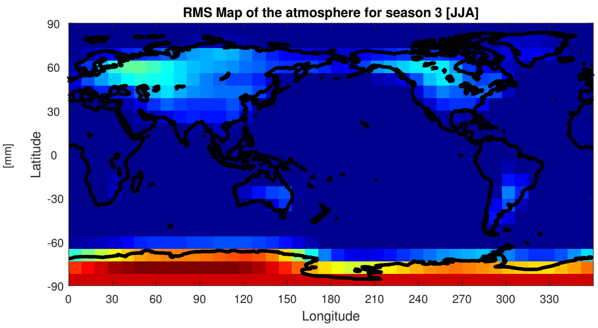

(c) RMS map from the daily solutions for season two



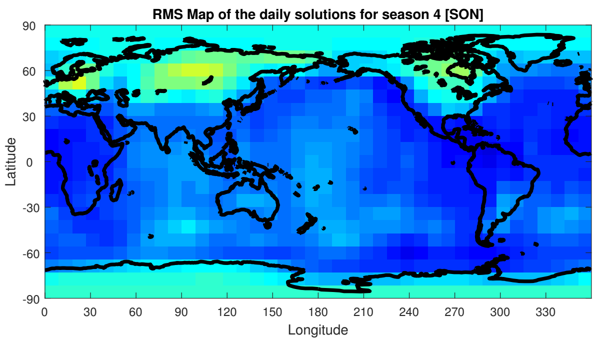

(d) RMS map from the atmospheric gravity fields for season two



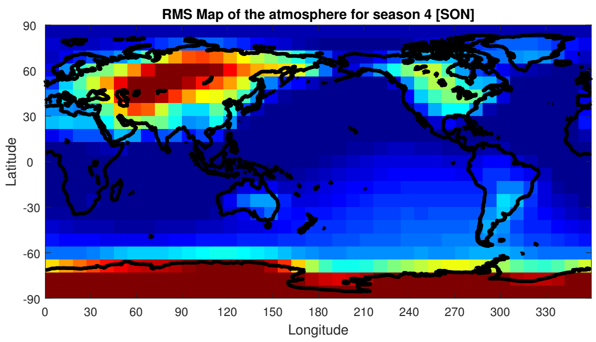

(e) RMS map from the daily solutions for season three




(f) RMS map from the atmospheric gravity fields for season three




(g) RMS map from the daily solutions for season four



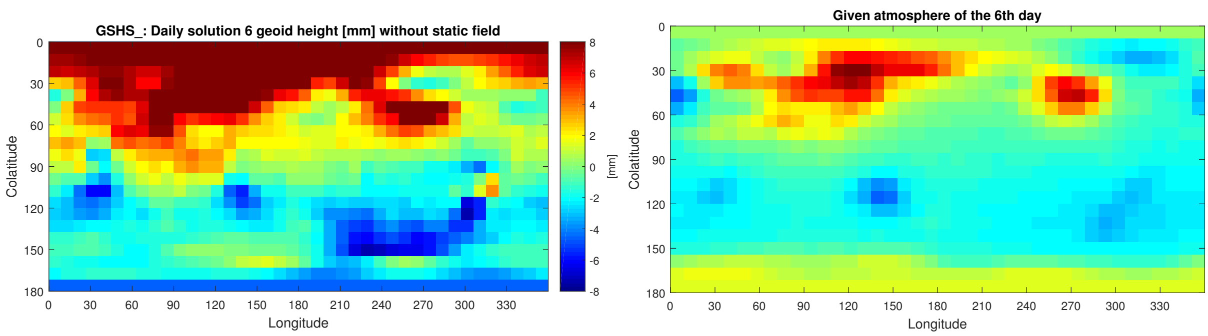

(h) RMS map from the atmospheric gravity fields for season four

***Figure 2.19:*** *RMS maps of the seasons*

In figure 2.19 it can be seen that there are also many similar patterns, as we already saw in the RMS maps of the whole year in the figure 2.18. But in these comparisons we can see the differences in the intensities of the pattern much better. The pattern from the daily solutions are similar to the pattern of the atmospheric gravity fields but very weak, except the pattern over North America of the first season. The third season from the daily solutions shows also less similarities with the gravity field of the atmosphere in the northern hemisphere, which can be explained with the weak positive pattern of the atmospheric gravity field in this area, because in the other seasons, only the strong pattern of the atmospheric gravity field match with the daily solution. Again, with this comparison it can be assumed that there are indeed some similarities, which is a good indicator that we can find significant signals from the gravity fields of the atmosphere in the daily solutions. After we have seen the first comparison by using the RMS map, we can now start with further comparisons in the next section.

## 2.7 Comparison of the daily solutions and the atmosphere

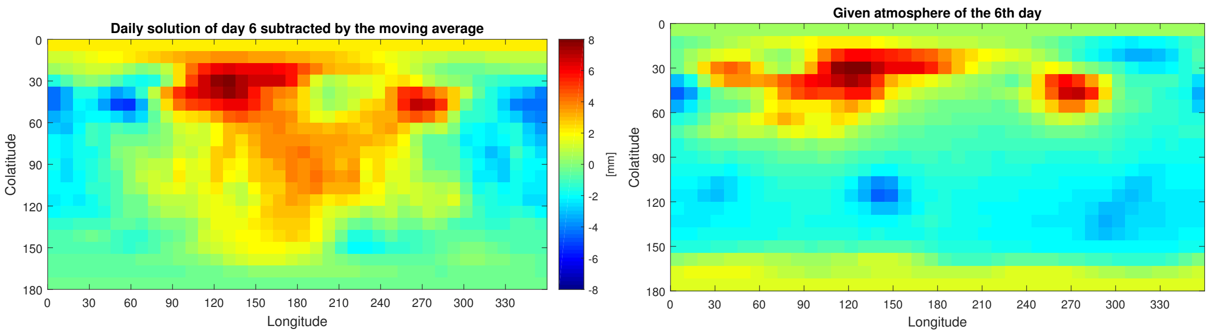
Now, we can compare the given model of the atmosphere and the simulated measured gravity field, given by the daily solutions. So in other words, we compare now two gravity fields with each other. In the following, we compare in a first step the 6th day of the month, to get a first idea, how closely these fields are related to each other. As already mentioned in the previous chapter, it is possible to subtract the static field from the daily solution, to eliminate the dominating ellipsoid, which covers the small variations we want to see. But it is also possible to subtract the moving average field, to eliminate also longer term variations which cover the small and short term variations from the atmosphere, which we want to reveal, as it could be seen in figure 2.16 there are even in 6 hours step many variations. For the sake of clarity, the term atmosphere will be used instead of "atmospheric gravity fields" and "gravity fields of the atmosphere" starting with this chapter. To show the effects of them, we have at first a comparison of the first day of the daily solution and the atmosphere, where only the static field is subtracted from the daily solution.




(a) Daily solution of the 6th day (b) Atmosphere of the 6th day

***Figure 2.20:*** *Comparison between the daily solution of the 6th day which is subtracted with the static field and the atmosphere of the 6th day until degree 20*

It is already possible to see similarities between the atmospheric field and the daily solution, so some areas show the same pattern. But it seems that the signals are much stronger in the daily solution, so there are similarities in the pattern but not in the strength of the signals. Also, in the north of the daily solution map is a very strong positive field which does not appear in the atmosphere. In the opposite, the atmosphere has a weak positive field in the south, where the daily solution has a negative field. So, there are several points to improve, if we want a higher correlation between the daily solutions and the atmosphere. In the next step, we will reduce the daily solution with the eleven day mean field to see the benefit by comparing it with the atmosphere:

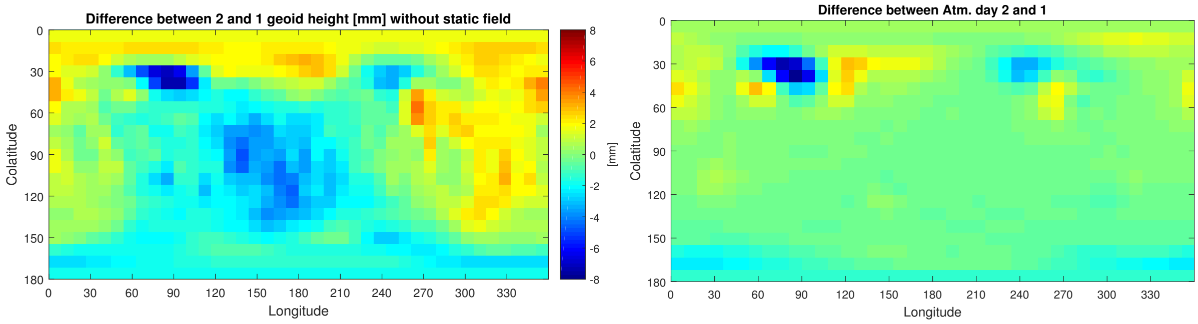



(a) Daily solution of the first day subtracted by the moving average

(b) Atmosphere of the 6th day

***Figure 2.21:*** *Comparison between the daily solution of the 6th day which is subtracted with the static field and also with the moving average and the atmosphere of the 6th day until degree 20*

If we have now a look at these images in figure 2.21, it becomes obvious that here are now much bigger similarities between the two images and we can see that more similar pattern occur, after the eleven day mean field is subtracted. But there are also patterns which do not match with the patterns from the atmosphere, which can be caused by other effects, like hydrology or the ocean. Beside that, we already discovered that there is a big improvement by subtracting the eleven day mean field. Another thought during the work with the daily solution was, to compare the changes between two days, because regarding the daily solutions, we could see big changes by calculating the difference. This will be done now, by comparing the changes between the first and the second day on the one hand for the daily solutions, and to the other hand for the atmospheric data:




(a) Differences between the daily solution of the second and first day

(b) Differences between the Atmospheric field of the second and first day

***Figure 2.22:*** *Here can be seen the comparison of the result by subtract the daily solution from the second day and the first day and the subtraction of the atmospheric field from the second day and the first day*

These images in figure 2.22 represent the best result, in which part of the atmospheric field clearly can be seen in the daily solution. The reason for this improvement could be the subtraction of two days, all effects which change in a low frequency are eliminated by the subtraction. In the opposite, the atmosphere change very fast, which could be seen very well, with the plotting of the first 48 hours in 6 hour steps. So these fast changes occur in the subtraction of the two days. It can be said that the effects of the atmosphere can be seen in the daily solutions. So we reached a big milestone in our goal, to estimate the atmosphere, by estimating signals from the atmosphere in the daily solutions.

## 2.8 Using the correlation coefficients to estimate the accordance

In the last section, we only compared the daily solutions and the atmosphere with our eyes, which is for a first test quite good, but to make a more significant analyse, we have to search for possibilities to compare the atmosphere and the daily solution in a mathematical way. This can be done by calculating the correlation coefficient $r$ for the daily solutions and the atmospheres. This value shows the linear correlation between two matrices. It can reach values between -1 to 1 that means when it has the value 1 (or -1) the two matrices have full correlation (or full negative correlation). It is to mention that $r$ has no dimension. The formula to calculate this is (in the MATLAB function 'corr2') the following:

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{(\sum_m \sum_n (A_{mn} - \bar{A})^2)\,(\sum_m \sum_n (B_{mn} - \bar{B})^2)}} \tag{2.24}$$

In this formula 2.24, $\bar{A}$ and $\bar{B}$ are the mean values from A and B. When we calculate this for each of the pairs of matrices, which we get when we calculate the difference between two following days, the results can be seen in the following graphic. It should be mentioned that the moving average field is subtracted by from the daily solutions:

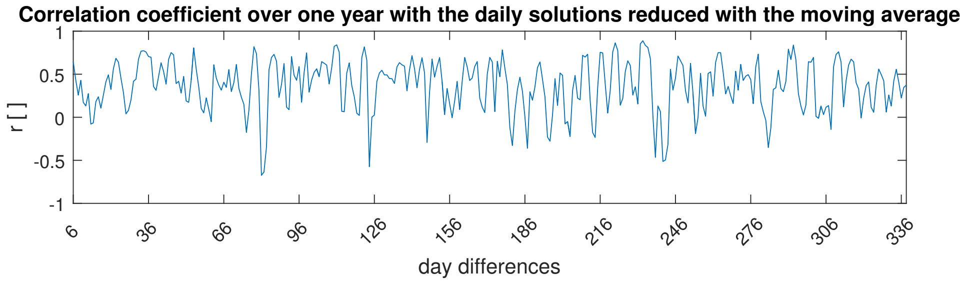



***Figure 2.23:*** *The correlation coefficient over one year*

Here, a very mixed result can be seen. We have some very good results in the correlation, but at the same time also several bad ones. If we have a look at the numbers, there are 130 days, in which the correlation is higher than 0.5, which is by a sum of 353 days, not even half of the measured days and the mean value is with 0.3690 also very bad. So we have some correlations but there is no significant amount of them to state that there are significant signals from the atmosphere in the gravity field of the earth. Until now, during the calculation of the correlation coefficients, we used only the potential fields from the daily solution and the atmosphere as planes, but if we multiply each row with the corresponding sinus theta term, which is normalised, we can calculate the correlation on a sphere:

$$V_{pot_{sphere}} = s\bar{i}n\theta \cdot V_{pot} \qquad \text{with } s\bar{i}n\theta \text{ as the normalized } sin\theta \tag{2.25}$$

Using this, we receive the following correlation coefficients over one year:

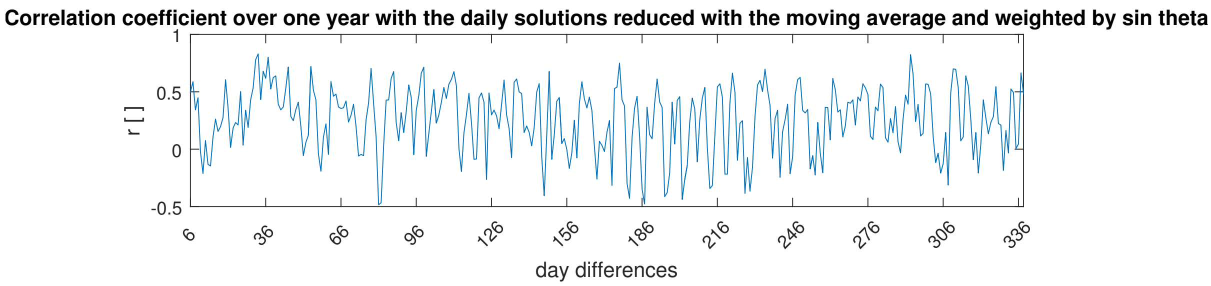
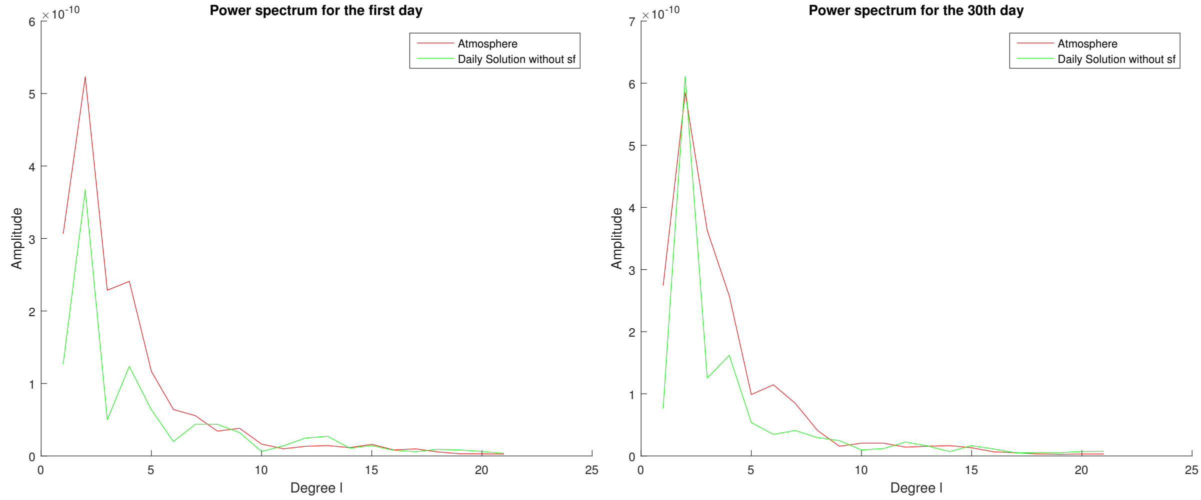



***Figure 2.24:*** *The correlation coefficient over one year weighted with sin theta*

But a look at this graph in figure 2.24 shows that the result looks worse than the previous one, also a look at the numbers shows us a worse result, where 73 are higher than 0.5 and the mean value is 0.2570 which is no really good results. All in all, it can be said that these results are not very motivating. But if we remember that we only test if in the unchanged daily solutions are signals from the atmosphere it is not that bad, because there is also a big room to improve the result. So, in further tests we have to think about procedures, to make the signals more visible. In another step, we can also compare the daily solutions with the atmospheric data in the power spectrum, to see, if we can see there some similarities between them. The problem, which occurs when we analyse the spectrum is that if there are no similarities between the two spectrum from the atmosphere and the daily solution, it does not mean, that they are not correlated with each other. The formula to calculate it is:

$$\sigma_l = \sqrt{\sum_{m=0}^{l} s_{l,m}^2 + c_{l,m}^2} \tag{2.26}$$

The amplitude of the graphics is caused by the summation of the spherical harmonic coefficients in the square. The graphics of the power spectrum of the atmosphere and the daily solution without the static field of the first and the 30th day can be seen below:




(a) Power spectrum of the atmosphere and the daily solution for the first day

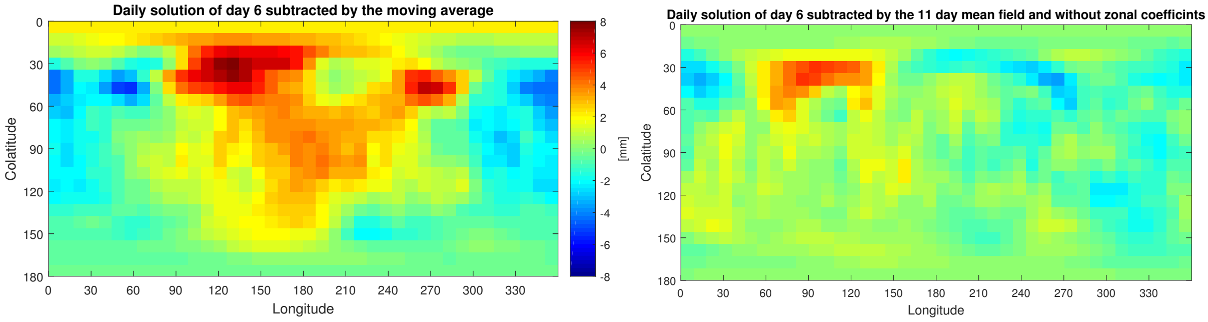
(b) Power spectrum of the atmosphere and the daily solution of the 30th day

***Figure 2.25:*** *Power spectrum of the atmosphere and the daily solution*

For the first day on the left side of figure 2.25, we can see some similarities, in which the amplitudes of the daily solution is less than the amplitudes of the atmosphere. If we have a look at the 30th day in figure 2.25, no clear similarities can be seen. As said before, this does not mean that they are not correlated, so it can be said that this method is not useful to compare the atmosphere and the daily solution. But as we could see in the images, there are for example in the polar region horizontal line pattern, which could be caused by measurement errors. We can try now to test, if there is an improvement, if we eliminate them.

## 2.9 Re-movement of the zonal coefficients

As already mentioned, we were able to see belts in the images of the daily solutions, mostly located in the polar regions, which could have been created during the measurement. If this is true and we have an error source in this region, it could help to eliminate the zonal coefficients that are coefficients which have the order m = 0, so parts of the signals are eliminated now. To see the differences, the next comparison shows the 6th day of the daily solutions subtracted by the moving average field with zonal coefficients and the same field without zonal coefficients:

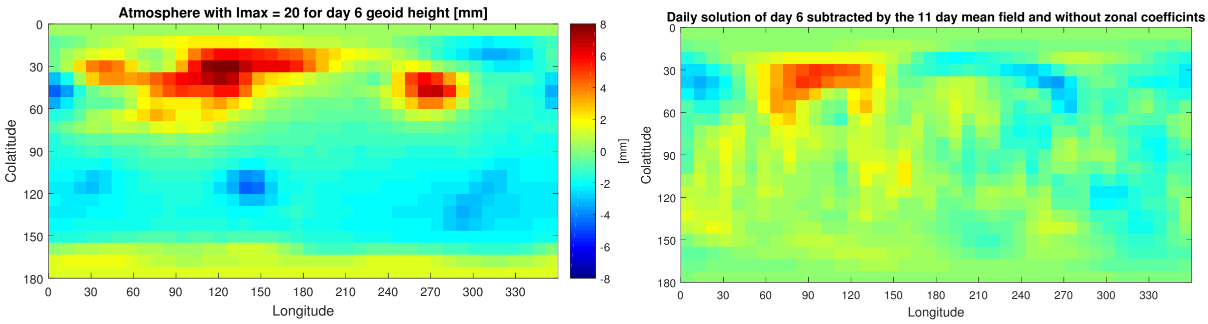



(a) Daily solution of the 6th day subtracted by the moving average

(b) Daily solution of the 6th day subtracted by the moving average without the zonal coefficients

***Figure 2.26:*** *Daily solution subtracted by the moving average with and without zonal coefficients*

Figure 2.26 clearly shows that after the subtraction of the zonal coefficients, the belts disappear and more fine structures can be seen, which were covered by the structures, created from the zonal coefficients. Now it can be tested, if this is also helpful for the comparison with the atmosphere, which can be seen in the next pair:




(a) Atmosphere of the 6th day

(b) Daily solution of the first day subtracted by the moving average without the zonal coefficients

***Figure 2.27:*** *Atmosphere and the daily solution without zonal coefficients of the 6th day*

This comparison in figure 2.27 shows that there is indeed some improvement, because in the equatorial area of the atmosphere, there are very weak signals while in the previous version of the daily solutions, which contained the zonal coefficients, a strong positive pattern in this area was visible. After the subtraction of the coefficients, this pattern still exists, but it decreased, which creates a higher similarity with the atmosphere. Now, we can calculate the correlation coefficients to see, if the changes can be also seen here. In this case, the coefficients over one year are calculated:

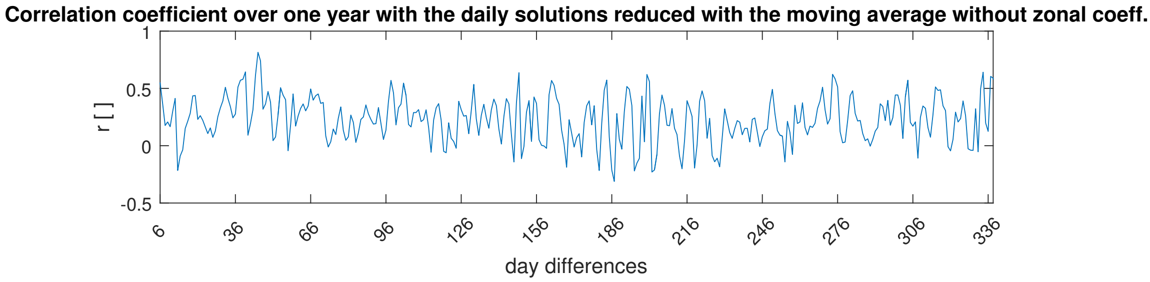



***Figure 2.28:*** *The correlation coefficient over one year without zonal coefficients*

In this case, 29 coefficients are higher than 0.5, which is a very low number and shows a worse result than the previous analysis. Now, it is very interesting to see, which result the using of the $sin\theta$ multiplication has, which was described in the section before:

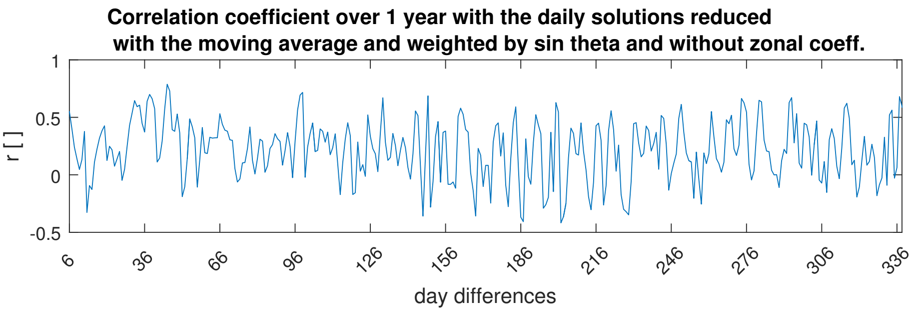



***Figure 2.29:*** *The correlation coefficient over one year without zonal coefficients and by using the multiplication of sin theta*

In figure 2.29 are 52 coefficients over 0.5, which is still low, but an improvement to the previous case. The highest number is 0.7876 at day 342, also 52 coefficients are lower then 0. All in all, it can be said that the subtraction of the zonal coefficients is on the one hand an improvement to see small structures, which are covered in the other case by the belts, which the zonal coefficients create. On the other hand, if we have a look at the correlation coefficients, it is no improvement, because the number of elements is very low. There is still one point, which we can test, if we remember that we have for each day four atmospheric epochs, we can also test, in which way each of them match with the daily solution.

## 2.10 Comparison of the atmospheric epochs and the daily solution

Because there are 4 epochs of atmospheric data for every day (0 - 6h, 6 - 12h, 12 - 18h, 18 - 24h) the question occurs, if every atmospheric epoch has the same contribution to the atmospheric signal, which is contained in the daily solution or are there differences in their contribution. For this, we compare for each epoch the correlation coefficients in several ways, with the normal correlation and by weighing the potentials with $sin\theta$, then also both with the subtraction of the zonal coefficients. Because there would be a lot of graphs in this section and if we also have a look at all mean values and other aspects, it will become confusing in a short time. Thus, all figures are listed at the end in the Appendix A.1, while below, only the table in which all results are listed will be shown. Now we have a look at the table to see, if there is an improvement in the comparison of the daily solution with the epochs of the atmosphere:

***Table 2.1:*** *Comparison of the epochs of the atmosphere with the daily solution*

| Atmospheric epoc | 0 - 6 h | 6 - 12 h | 12 - 18 h | 18 - 24 h | 12 - 24 h |
|---|---|---|---|---|---|
| Number of coefficients higher 0.5 | 115 | 131 | 130 | 130 | 131 |
| Mean coefficient | 0.3367 | 0.3654 | 0.3779 | 0.3766 | 0.3789 |
| Number of coefficients higher 0.5 with $sin\theta$ | 57 | 74 | 80 | 84 | 88 |
| Mean coefficient with $sin\theta$ | 0.2291 | 0.2542 | 0.2646 | 0.2573 | 0.2633 |
| Number of coefficients after removing the zonal coefficients | 41 | 26 | 31 | 34 | 34 |
| Mean coefficient after removing the zonal coefficients | 0.1951 | 0.2284 | 0.2378 | 0.2356 | 0.2378 |
| Number of coefficients after removing the zonal coefficients, with $sin\theta$ | 41 | 49 | 51 | 58 | 53 |
| Mean coefficient after removing the zonal coefficients, with $sin\theta$ | 0.1951 | 0.2177 | 0.2265 | 0.2193 | 0.2252 |

It can be said that to compare each epoch of the atmosphere with the daily solution is a useful method. It shows that the daily solution indeed contains mostly only information from the last two epochs. While the results from the unchanged fields show nearly similar results, except the first epoch, the real improvement by using the third and fourth season can be seen by using $sin\theta$. Interestingly, the summation of both showed only mixed results, in some calculations it shows high values while in other comparisons it is lower or equal to the single epochs. The fourth epoch has in the first correlation versions the same number of coefficients over 0.5 than the third but in the combination of both the value is a bit higher and also the mean coefficient is higher. It has to be mentioned that beside all, the number of correlation coefficients over 0.5 is very low, if we regard the fact that the total number of coefficients is 353, the best result over 0.5 is with 131 not even the half of it but compared to the correlation with the whole atmosphere which is 130, we see that this is equal to the results of the third and fourth epoch and with the combination, we can improve it a little.

After this chapter, we have a good idea, how the data we use is created and also how the daily solutions are correlated to the atmosphere. In the next chapter we will have a look at the possibilities to estimate the atmosphere by using the daily solutions.

# Chapter 3

# Estimating the atmosphere

## 3.1 Overview

Now, after we learned all about the data in the second chapter and already made some comparisons between the atmosphere and the daily solutions, we are well prepared to start with the estimation of the atmosphere. Doing this, we will use only the given data to examine the atmosphere. Thus, the given atmospheric fields are subtracted from the daily solutions to create *masks*, which will be used in a next step to estimate the atmosphere of each day. This will be done by subtracting the *masks* from the daily solutions. For the estimation of the atmosphere, the mean fields of the differences between the daily solutions and the atmosphere can cover eleven days or the whole month. After this, we test the possibility to use the differences of the daily solutions to calculate the whole atmosphere for one year. This method needs an atmospheric model only as a 'starting day', all other atmospheric days are calculated again by only using the daily solutions. At least, it has to be said that all daily solutions which are used in the calculations are subtracted by the static field and also from the moving average field. Only in special cases, the daily solution is used without the subtraction of the mean field, but in these cases it is said explicitly. For both ways to estimate the atmosphere we will make several tests, in order to compare it with the given atmosphere. In a first step, we will have a look at the RMS maps to see, if there are similar energy patterns. After that, we calculate the correlation coefficients, which give us an impression how similar the estimated and given atmosphere are. Then, we will look at the power spectrum and, beside that, also analyse, if there is a factor difference between the given atmosphere and the estimated atmosphere. At least there will be a short look into the frequency domain, to see, if there are also similarities. With these tests, we will have a good impression, how good the new calculated atmosphere matches with the given atmosphere. But before we really start to estimate the atmosphere, we have to analyse one important point. To make our solutions reasonable, we have to figure out, in which way the used data for the estimation is related to each other. This will be tested now in a first test, before we start with the estimation of the atmosphere.

## 3.2 Covariance function

In a first test, the differences between the atmosphere and the daily solutions, which are subtracted from the static field and the given moving average fields, will be calculated for a whole year, which will be also represent basic data to calculate the *masks*:

$$diff_i = atm_i - ds_i \tag{3.1}$$

Afterwards, the covariance will be calculated between these differences:

$$cov_i = cov(diff_i, diff_j), \qquad |i - j| = const \tag{3.2}$$

Where $i$ is the 'starting day' and $j$ is the number of days between the 'starting day' and the actual day. By using the MATLAB function $cov()$ on the matrices of the potential, the covariance matrix is calculated:

$$cov(\mathbf{A}, \mathbf{B}) = \begin{pmatrix} cov(\mathbf{A}, \mathbf{A}) & \underline{cov(\mathbf{A}, \mathbf{B})} \\ cov(\mathbf{B}, \mathbf{A}) & cov(\mathbf{B}, \mathbf{B}) \end{pmatrix} \tag{3.3}$$

Where **A**, **B** are matrices of the potentials from the atmosphere and the daily solutions. Thus, to get the covariance of the matrices **A** and **B**, the underlined upper right element has to be taken. While the constant difference between the days will be increased, the mean value of all covariance coefficients will be calculated for each step, to compare it more precisely to the other calculated values. If we plot now these mean covariance coefficients, we receive the following graphic:

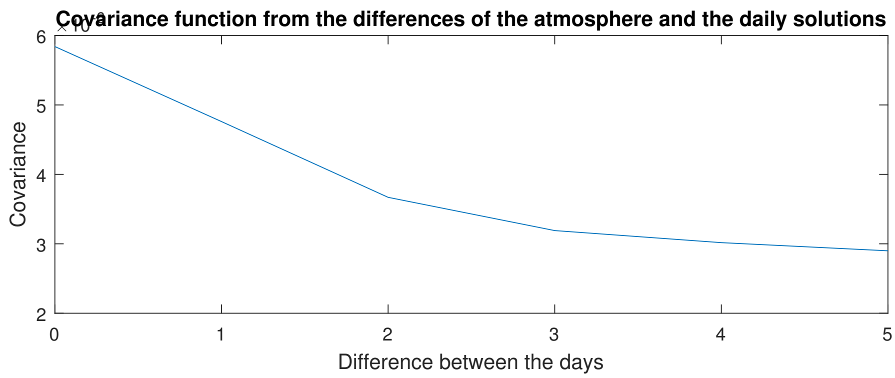



***Figure 3.1:*** *Calculated covariance function for daily solutions*

As it can be seen here, the graph has a slow decrease. That means that the calculated differences are related to each other, so it is possible to use them for the estimation, because if we want to use them to create our *masks*, we make it with the assumption that the fields are correlated with each other. With this test, we have the assurance that this is the case. In the next step, it is also possible to calculate the covariance function of the differences between the daily solutions, which is the basic data for the other atmospheric estimation:

$$\begin{gathered} diff_i = ds_{i+1} - ds_i \\ cov_i = cov(diff_i, diff_j) \quad |i - j| = const \end{gathered} \tag{3.4}$$

Like in earlier calculations, the mean value will be calculated by using these values. If we plot them, the result can be seen in the following:

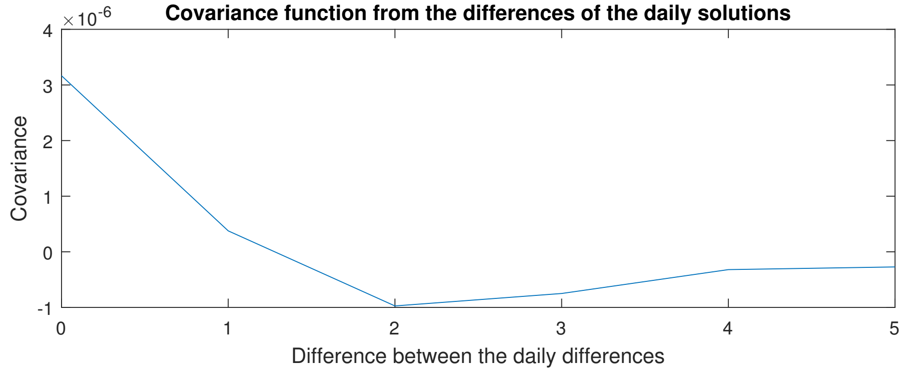



***Figure 3.2:*** *Calculated covariance function for the differences between daily solutions*

This result is worse than the previous one, because the graph decreases very fast, which means, the difference of one day has not much similarity to the difference of the daily solutions, which is from two days later. The problem with the graphics before is that the y-axis is in a scale, which is not normalised and therefore difficult to interpret. Thus, it is helpful to use the correlation coefficient. The reason can be seen, when we have a look at the calculation:

$$Corr_i = \frac{\sum_i (A - \bar{A}) \cdot (B - \bar{B})}{\sqrt{\sum_i (A - \bar{A})^2 \cdot (B - \bar{B})^2}} \tag{3.5}$$

This can be written according to (*Wikipedia: Korrelationskoeffizient*, n.d.):

$$Corr_i = \frac{Cov(A, B)}{\sigma(A) \cdot \sigma(B)} \tag{3.6}$$

Where $\sigma$ is the standard deviation of the values $A$ and $B$. Thus, it is better to interpret the graphs, because now they are normalised between -1 and 1. If we use this, the graph of the differences between the atmosphere and the daily solutions looks like the following:

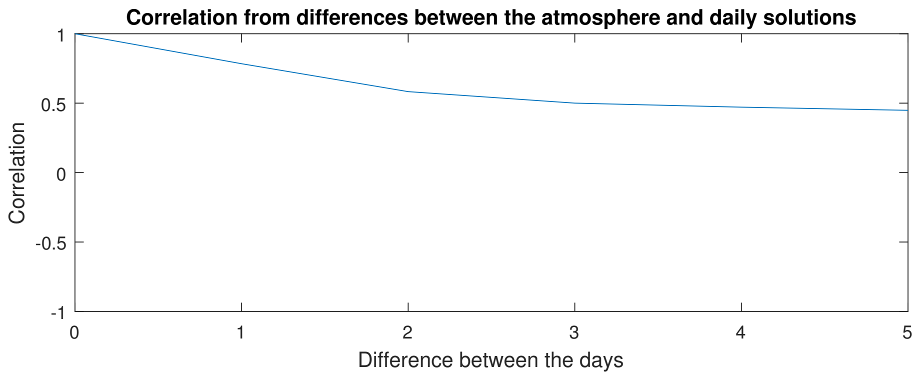



***Figure 3.3:*** *Calculated correlation for daily solutions*

Here, the high correlation between the differences can be seen clearly, even though there are 5 days between them. In contrast, if we have a look at the graph from the differences of two daily solutions, a worse result can be seen:

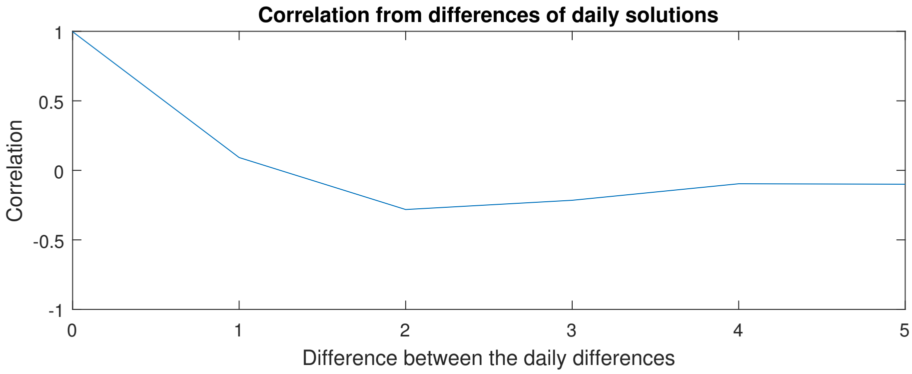



***Figure 3.4:*** *Calculated correlation for the differences between daily solutions*

The graph is after one day difference already nearly zero, so the daily differences of one day is not correlated with the differences after two days.

### 3.2.1 Conclusions of the results

As already said, the covariance of the differences between the atmosphere and the daily solution shows a very good result, regarding the use of them to estimate the atmosphere. Because they have a lot in common, it means they can be used very well to create *masks*, thus there is a big connection between the fields. In contrast, the covariance of the differences of the daily solutions is very low. That means that it is difficult to make predictions because the fields have no connections between each other after one day. Thus, if they are used to estimate the atmosphere, the correlation of the estimation with the real atmosphere depends only on the actual field of the differences. There is no connection to previous fields. As a conclusion, it can be said that it is promising to use the differences between the atmosphere and the daily solutions to estimate the atmosphere, instead of using the differences of the daily solutions. Therefore, we will focus on creating the *masks*, but later we will also have a look at the other possibility, how the low correlation affects the estimation of the atmosphere.

## 3.3 Estimating the atmosphere by calculating *masks*

Now, after we have an idea about the correlation between the used data, we can start to create the *masks* for the atmospheric extraction by using eleven day means of the differences and the monthly mean of them with the following calculation:

$$field_{diff_i} = ds_i - atm_i \tag{3.7}$$

With:

$field_{diff}$ = field of the differences
$ds$ = daily solution which is subtracted by the static field and the moving average
$atm$ = atmosphere
$i$ = actual day which is calculated

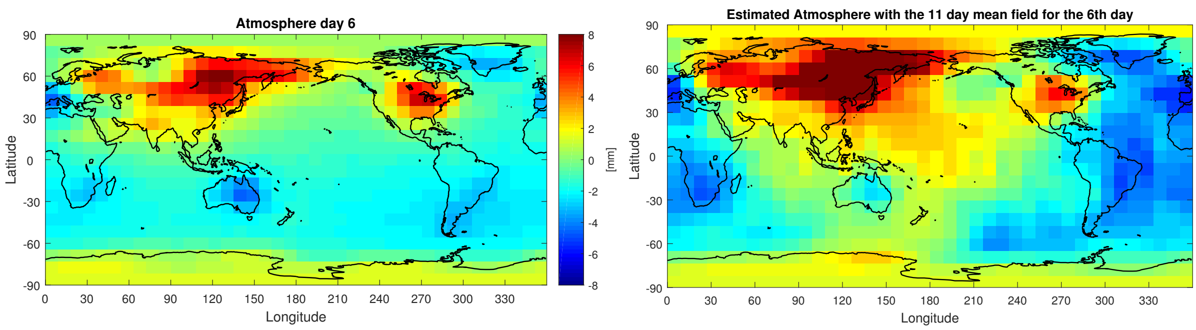
So it is to mention, that in this calculation the daily solutions are subtracted by the moving average. After that, the *masks* can be calculated:

$$mask_j = \frac{1}{k} \cdot \sum_{i=1}^{j*k} field_{diff,i} \tag{3.8}$$

Where $k$ is eleven if we want to calculate the 11 day mask, or it is the number of the days of the month, if we want to calculate the monthly masks. $j$ is the number of the mask, which is actually calculated. With that, the new atmosphere can be calculated now in this way:

$$atm_{new,i} = ds_i - mask_j \tag{3.9}$$

Here, $i$ is the actual day, which is calculated and $j$ is the corresponding eleven day period which covers the actual day. Because in the eleven days are no big variations, compared to the monthly results, it can be expected that there will be a higher correlation with them. But, considering the fact that one goal of the work is to create a method also for other years, it is also important, that the monthly mean field shows good results, because the atmosphere has periodic changes and also atmospheric models for each month exist. Thus, it is already possible to extract the atmosphere for each month. Because the atmosphere changes periodically, it would be possible to create standard *masks* for each month, to extract the atmospheric signals. Now, we want to have a first look at the extracted atmosphere in comparison of the real atmosphere:

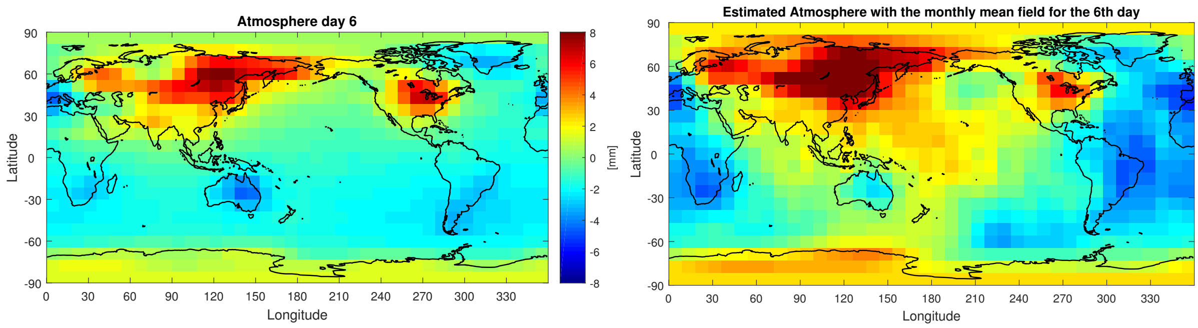



(a) Atmosphere of the 6th day

(b) Estimated atmosphere of the 6th day by using the 11 day mean from the differences

***Figure 3.5:*** *Comparison between the atmosphere of the 6th day and the estimated atmosphere by using the 11 day mean field*

It can be seen in figure 3.5 that there is a high correlations between the atmosphere of the 6th day and also the estimated atmosphere, which is calculated by the eleven day mean mask. But there are two big differences, the positive area over China in the estimated atmosphere is much bigger than the positive area in the same region in the given atmosphere. Also over the Pacific is a weak positive area in the estimated atmosphere, which can not be seen in the given atmosphere. At least, there are several stronger negative pattern in the estimated atmosphere, where the given atmosphere show only weak negative pattern. However, it can be said that all striking pattern from the atmosphere can be seen also in the estimated field. Now, we can have a look at the comparison of the same day of the atmosphere compared with the estimated atmosphere by using the monthly mean field. Because there are more atmospheric fields included by the calculation of the mean field, it can be expected that there are more variations.




(a) Atmosphere of the 6th day day

(b) Estimated atmosphere of the 6th day by using the monthly mean from the differences

***Figure 3.6:*** *Comparison between the atmosphere of the first day and the estimated atmosphere by using the monthly mean field*

As already expected, more variations can be seen between the two fields in figure 3.6, but they are smaller than expected. Beside that, there are still some comparisons between the two fields and it is remarkable that the variations are similar to the previous ones. Again, there is a big positive pattern over the Pacific and a bigger negative pattern in the right area but here they are weaker then in the estimated atmosphere by using the 11 day masks and the striking areas from the atmosphere can be identified well. As a first conclusion, it can be said that there are not many differences in this example between the atmosphere, estimated of the eleven day mean field of the differences and the one, where the whole month is used. Both fields show many similarities with the atmosphere of the same day.

### 3.3.1 Regarding further comparisons

As done in the previous chapter, we start with a comparison by only using our eyes, which is good for the beginning to have a first impression how the estimated data look. After that, more analytical methods will be used to have a closer look at the correlation between the two estimated atmospheres and the given atmosphere. In a first step, the correlation coefficients will be calculated, which are already described in eq. 2.24:

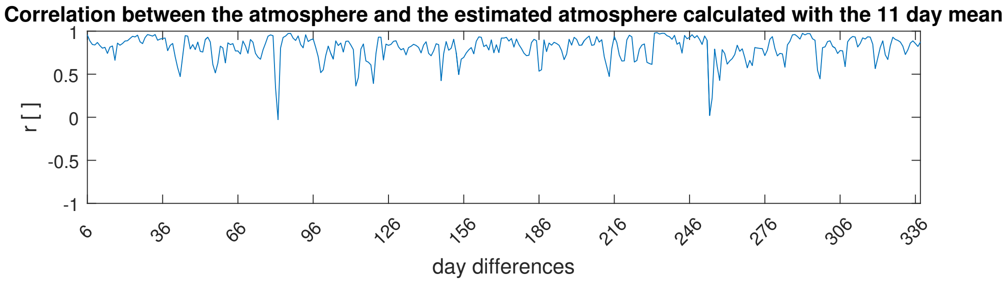



*Figure 3.7: The correlation coefficients over the whole year for the comparison between the atmosphere and the estimated atmosphere by using the 11 day mean field of differences*

In the graph in figure 3.7, 340 coefficients out of 353 are higher than 0.5 and 347 are higher than 0.4, which is a good result. The mean coefficient is 0.8079, which is also a good number. Now, also the correlation, by regarding the spherical form of the earth can be calculated, by multiplying the fields with the $sin\theta$ values. In this case, the following results can be seen:

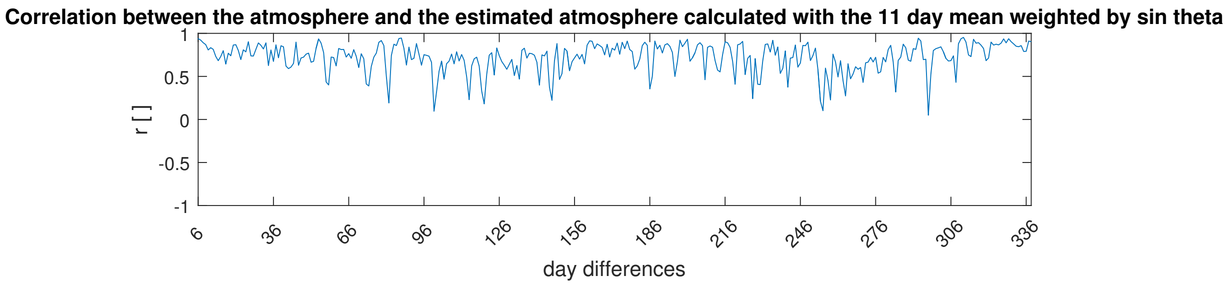



*Figure 3.8: The correlation coefficients over the whole year for the comparison between the atmosphere and the estimated atmosphere by using the 11 day mean field of differences, by using the $sin\theta$ values*

In this case, 316 coefficients are higher than 0.5 and 334 higher than 0.4, which is a bit lower than the coefficients, which are calculated without the $sin\theta$ values, but it is still a good result. The mean coefficient is in this case 0.7223, which is again lower as the previous one, but still good. Now the same comparison can be made with the estimated atmosphere, which is calculated by using the monthly mean field of the differences. At first, the normal correlation is shown:

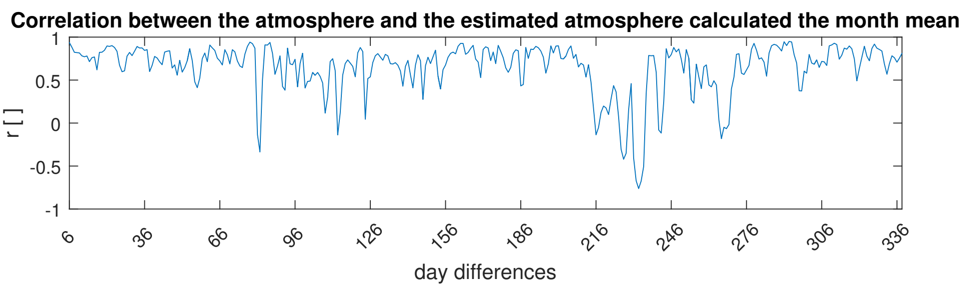



*Figure 3.9: The correlation coefficients over the whole year for the comparison between the atmosphere and the estimated atmosphere by using the monthly mean field of differences*

A look at the figure 3.9 shows a worse result than the previous one. Here, 288 coefficients are higher than 0.5 and 310 higher than 0.4, which is a high number but lower than the previous ones. However, it could be expected that in this case the fields show a higher variation compared to the given atmosphere, which can be seen here in the lower numbers of similar days. The mean value is 0.6529, which is lower compared to the previous ones but still good and above 0.5. In the next step, we can have a look at the correlation, where the fields are multiply with $sin\theta$:

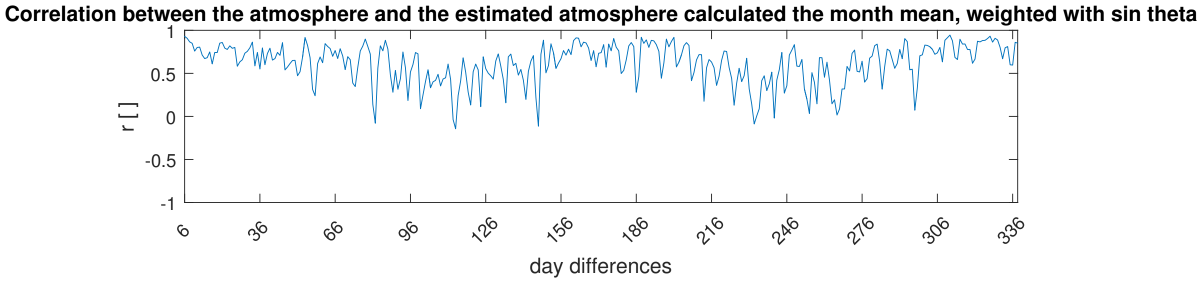



***Figure 3.10:*** *The correlation coefficients over the whole year for the comparison between the atmosphere and the estimated atmosphere by using the monthly mean field of differences, by using the sinθ values*

As it could be expected, the graph in figure 3.10 has 270 coefficients higher than 0.5 and therefore a lower value than in the previous analysis, where the spherical shape of the earth is not respected. If we look at the values over 0.4 we have with 300 still a good result. The mean coefficient is 0.6210 and it is surprisingly only a bit lower than the coefficient before. If we see these results, it can be said that there are high similarities between the estimated atmospheres and the given atmospheres. To test the impression of the different power of the coefficients, it is also possible to show graphics of the degree RMS, which show the RMS over the degree, by using the formula ( 2.26). Below, the degree RMS of the first day of the atmosphere, calculated by the eleven day mean differences is shown:

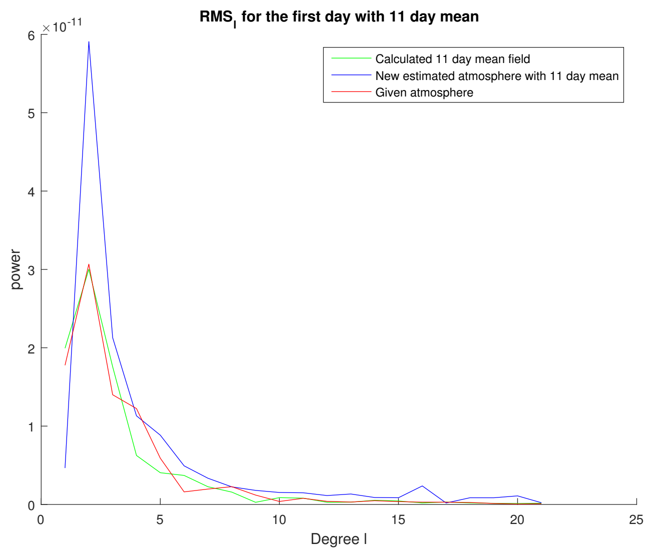



***Figure 3.11:*** *RMS plotted over the degree of the coefficients*

Figure 3.11 shows that the power of the lower coefficients are higher from the calculated atmosphere, than from the given atmosphere, which means that over all frequencies the estimated atmosphere shows a higher energy. The eleven day mean field has a similar energy compared to the given atmosphere, it decrease faster at the beginning, later it shows again a nearly similar energy. If we have a look at the given atmosphere, we see that the energy decreases fast and after degree 10 it is close to zero. Because the estimated atmosphere still contains a certain degree of energy this could be one source of errors, because the given atmosphere has in this area nearly no signal. That issue could be analysed closer in further studies. The same test can be done for the atmosphere of the first day, which was calculated by the monthly mean of the differences:

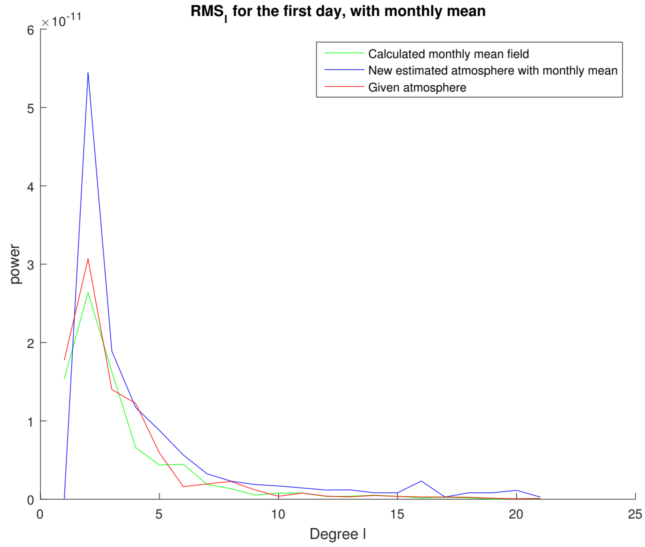
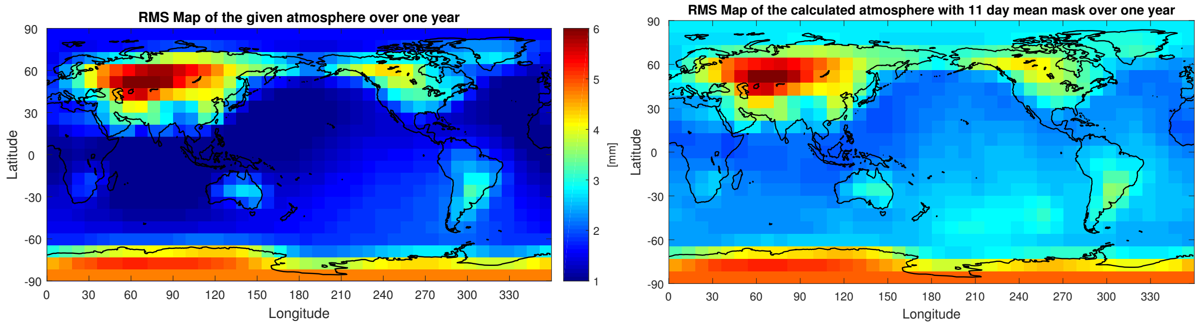



***Figure 3.12:*** *RMS plotted over the degree of the coefficients*

Again, a similar result as before can be seen, where the calculated coefficients are much stronger than the given one. But in this case, the monthly mean field shows at the beginning a lower energy compared to the given atmosphere, beside that, the curve of the estimated atmosphere shows many similarities to the curve of the estimated atmosphere by using 11 day mean fields. This shows again that there are no big variations between the atmosphere, calculated by the eleven day mean differences and the atmosphere, where the monthly means were used. In a last step, the RMS maps over one year can be calculated:

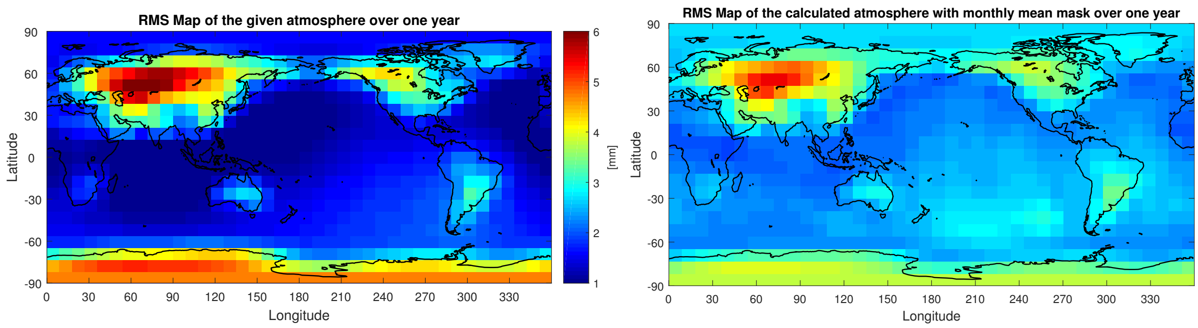



(a) RMS map of the given atmosphere over one year

(b) RMS map of the calculated atmosphere, by using the 11 day mean differences

***Figure 3.13:*** *Comparison between the RMS maps of the calculated atmosphere with 11 day mean fields and the given atmosphere*

As it can be seen here in the figure 3.13, there are big similarities between the calculated atmosphere and the given atmosphere. The main difference is that in the given atmosphere the patterns are stronger compared to the calculated atmosphere. Beside this, it can be said that the calculated atmosphere shows a high correlation with the given atmosphere. Now, we can also have a look at the calculated atmosphere with the monthly *masks*:




(a) RMS map of the given atmosphere over one year

(b) RMS map of the calculated atmosphere, by using the monthly mean differences

***Figure 3.14:*** *Comparison betweent the RMS maps of the calculated atmosphere with monthly mean fields and the given atmosphere*

Also here in figure 3.14 the similar pattern can be seen, besides the fact that there seems to be less energy in the RMS map of the estimated atmosphere. While in the previous comparison in figure 3.13, the calculated atmosphere shows only weak negative patterns and the positive pattern were stronger, in this case also the positive ones are weak. Because of that, it seems that the calculated atmosphere by using 11 day mean masks match better in some areas. For example in the positive pattern over China and east Russia can be seen better in the calculated atmosphere with the 11 day mean *masks* than the other one. Also the positive pattern over the Antarctica match better in the first comparison.

If we have a look at our results, it could be possible that the estimated atmosphere has a factor difference to the given atmosphere. To test this possibility, we can easily divide the calculated atmosphere with the given atmosphere to test the relation:

$$f_i = \frac{atm_{new,i}}{atm_i} \tag{3.10}$$

By plotting this, the following graph can be seen:

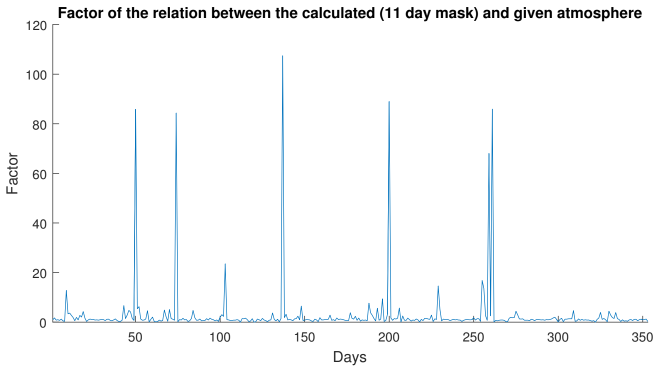
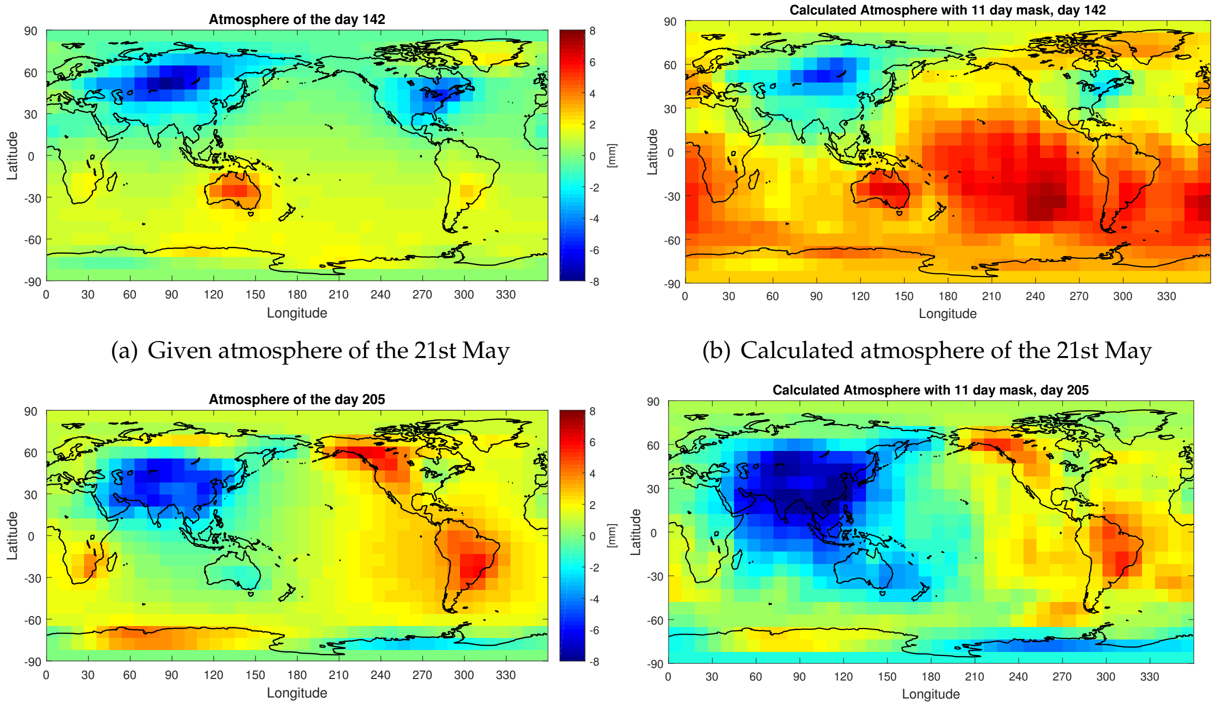



***Figure 3.15:*** *Factor between the calculated atmosphere and the given atmosphere*

This result in figure 3.19 is indeed surprising, because it could have been expected that there are some smaller variation, in truth, the variations go from $3.8543 \cdot 10^{-4}$ to 143.2259. All in all there are only 301 values lower as 2. If we remember the correlation coefficients (figure 3.8) we can not see similarities between lower correlation coefficients and the high peaks which can be seen in this figure 3.19. The biggest peak appear at the day 142, which is the 21st May. We already have seen in the RMS map 3.13 that we have the different strength in the pattern. Until now, we only looked at diagrams and numbers, but to have an idea, how the high peaks are created, it is helpful to have a look at the fields of the two days 142 and 205:




(a) Given atmosphere of the 21st May

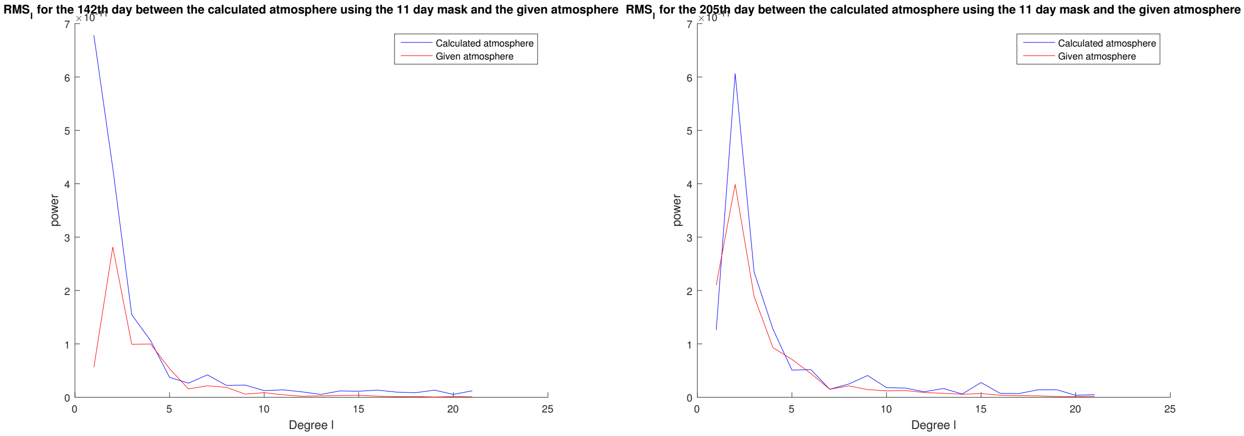
(b) Calculated atmosphere of the 21st May

(c) Given atmosphere of the 23rd July

(d) Calculated atmosphere of the 23rd July

***Figure 3.16:*** *Comparison between the given atmosphere and the estimated atmosphere*

If we have a look at the pattern in the two comparisons in figure 3.16, we get already an idea, why there is such a high peak and at the same time the correlation is still good. The patterns have a lot in common, which explains the high correlation, but the correlation does not respect the strength of the patterns, which lead to the high peaks. It clearly can be seen in this figure that at the 21st May, the calculated atmosphere is much stronger than the given one. At the 23rd July we have very strong negative patterns in the calculated atmosphere, which lead to the high peak. Now, it is also very interesting, how these days look in the power spectrum of the coefficients:




(a) RMS plotted over the degree of the coefficients for day 142

(b) RMS plotted over the degree of the coefficients for day 205

***Figure 3.17:*** *Comparison between RMS over degree for the days 142 and 205*

In figure 3.17 we can see clearly that for the highest peak at day 142 the coefficients show also a very high value, even the characteristic ascending part of the graph is missing, which can be seen in the graph of the day 205 for the estimated atmosphere. But beside this, both estimated atmospheres have nearly the same maximum, which is clearly higher than the maximum of the given atmosphere. At the estimated atmosphere of the day 205 we have still some smaller peaks at degree 9 and 15, which can not be seen at day 142. Because of the fact that the low degree has much energy, it means that the low frequencies have a very high energy at this field and with an increase of the frequency, the energy decreases. This effect is stronger at day 142 than at day 205. For later analysis, a Fourier analysis could be very interesting, but would extend the size of this work. Now, we can also have a look at the factor for the new atmosphere, which is calculated by the monthly masks:

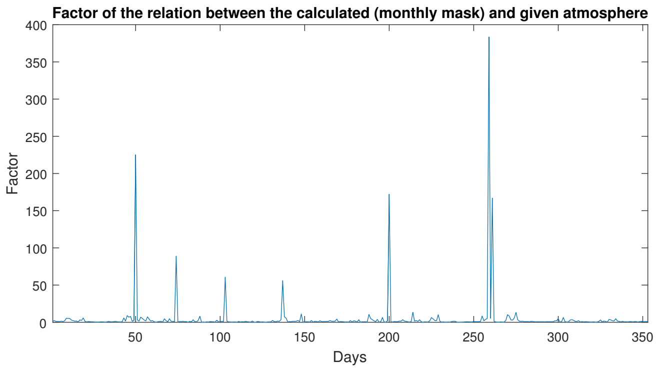



***Figure 3.18:*** *Factor between the calculated atmosphere and the given atmosphere*

As probably expected, there are more high peaks in figure 3.18 because of the fact that we now use a mean which is for a whole month. There are only 259 factors smaller then 2, so we have a significant increase of higher factors. We already see the peaks at nearly the same date only with a change in the strength. Now, with this knowledge, we can go forward as done before and compare the calculated atmosphere with the daily epochs of the atmosphere, in order to see if there are more correlations.

### 3.3.2 Comparison with the epochs

Because the test in the last chapter showed that the first two epochs have a lower correlation, we will focus here on the last two epochs. The analysis of the data will be done at the end of this section, to regard all results of the following analysis.

#### 3.3.2.1 Epoch 3

This epoch is from 12 - 18 o'clock:

***Table 3.1:*** *Estimated atmosphere compared with epoch 3*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 340 | 291 |
| Correlation coefficients $> 0.4$ [ ] | 347 | 309 |
| Mean correlation coefficient [ ] | 0.8082 | 0.6586 |

Regarding the spherical shape of the earth:

***Table 3.2:*** *Estimated atmosphere compared with epoch 3 regarding the shape of the earth*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 318 | 265 |
| Correlation coefficients $> 0.4$ [ ] | 331 | 296 |
| Mean correlation coefficient [ ] | 0.7174 | 0.6219 |

#### 3.3.2.2 Epoch 4

This epoch is from 18 - 24 o'clock:

***Table 3.3:*** *Estimated atmosphere compared with epoch 4*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 341 | 285 |
| Correlation coefficients $> 0.4$ [ ] | 348 | 310 |
| Mean correlation coefficient [ ] | 0.8058 | 0.6562 |

As done before, the spherical shape of the earth is regarded in the following:

***Table 3.4:*** *Estimated atmosphere compared with epoch 4 regarding the shape of the earth*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 305 | 268 |
| Correlation coefficients $> 0.4$ [ ] | 323 | 293 |
| Mean correlation coefficient [ ] | 0.7119 | 0.6162 |

#### 3.3.2.3 Epoch 3 and 4

This epoch is from 12 - 24 o'clock:

***Table 3.5:*** *Estimated atmosphere compared with epochs 3 and 4*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 342 | 287 |
| Correlation coefficients $> 0.4$ [ ] | 349 | 310 |
| Mean correlation coefficient [ ] | 0.8111 | 0.6606 |

At least the spherical shape of the earth is regarded for this epoch:

***Table 3.6:*** *Estimated atmosphere compared with epochs 3 and 4 regarding the shape of the earth*

| | 11 day mean masks | monthly mean masks |
|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 312 | 272 |
| Correlation coefficients $> 0.4$ [ ] | 330 | 295 |
| Mean correlation coefficient [ ] | 0.7211 | 0.6245 |

#### 3.3.2.4 Short conclusion of the comparison with the epochs

It can be said that the highest correlation is with the sum of the epochs 3 and 4, in this case, we have with the eleven day mean masks 349 higher than 0.4, which means that only 4 days have a lower correlation than 0.4. Also, there are 342 coefficients higher than 0.5, which is with 2 coefficients higher compared with the earlier comparison. With the monthly mean, we have a high correlation, surprisingly the correlation is, for this kind of estimation, in all three cases nearly the same with 309 - 310 higher than 0.4.

### 3.3.3 Conclusion for the differences of the daily solutions and the atmosphere

As a conclusion it can be said that this method shows very good results, with a number of 340 correlation coefficients which are higher than 0.5 out of 353. These results prove that the estimated atmosphere shows a very high correlation with the given atmosphere. We also can increase these results by considering only the last two epochs, where we have 349 coefficients higher than 0.4. Beside that, we found already a very good method to simulate the atmosphere by using the daily solutions, which can be used for further studies. One interesting point which should be analysed closer in later studies are the factor differences at special days, because closer analyse of this effect would exceed this work. But besides the good results, one problem could occur in this method. The point is that there could be atmospheric variations which are not included in the standard atmospheric model of each month. As said above, because of the global warming, more variations are possible, which perhaps are not included in the used *masks*, so that the wanted information could be lost or does not appear prominent in the estimated atmosphere. In this case, it would be helpful to have more information about the atmosphere and by using atmospheric models which include the latest variations of the atmosphere, the results could be improved considerably. This leads us to the next possibility to estimate the atmosphere, simply by adding the differences of two following daily solutions to each other.

## 3.4 Estimating the atmosphere for a whole year

### 3.4.1 A closer look at the differences of the daily solutions

It is already obvious that the differences between consecutive days of the daily solutions show high similarities compared to the differences between consecutive days in the atmosphere. It is to mention that in this section the used daily solutions are only subtracted by the static field, thus no moving average is needed and we have 363 days instead of the 353 days compared to the previous chapter. Now, before we start, we want to have a closer look at these similarities, by calculating the correlation coefficients for a whole year:

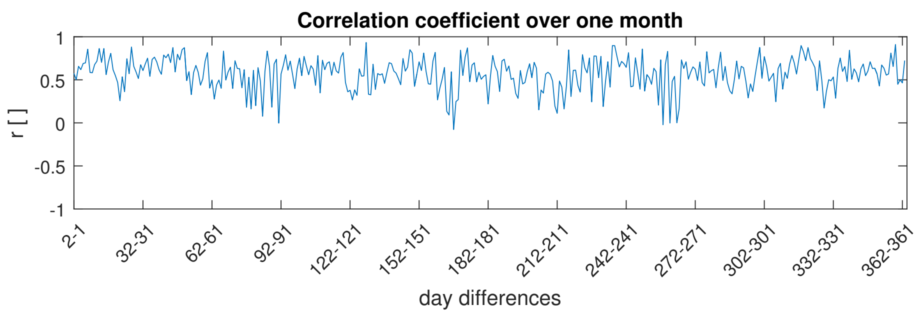



*Figure 3.19: The correlation coefficient over one year between the differences of two following days of the daily solution and of the atmosphere*

Here, out of 363 days, 108 are below the value 0.5 but out of these only 58 are below 0.4. So it can be said that from a whole year, 304 are higher than 0.4, which is a good result. That means otherwise that there is also a high correlation between the information of these differences. Thus, it is possible to create the atmosphere, by using the differences of the daily solutions. As a start field, which is needed to start with the estimation, we can use a standard atmosphere for January, in this case, the atmosphere of the first day is used. The advantage of this method is that it can be expected that at least the most variations of the atmosphere are included in the differences, so that they also appear in the estimated atmosphere. This leads to the benefit that only one atmospheric model is needed to estimate the atmosphere, and the estimation of the atmosphere does not depend on the availability of atmospheric models for the whole year.

### 3.4.2 Testing the correlation between the calculated and the given atmosphere

#### 3.4.2.1 Correlation with the whole atmosphere

In a first step, we can start to use the given first day of the atmosphere (1st January) and add on it the difference between two days of the daily solutions. At this point it is important to note that the daily solutions used in this method are only subtracted by the static field, because by calculating the differences only the changes between the fields appear, which are already short term variations. As shown before, these days show a high correlation with the differences of two consecutive days of the atmosphere. Before we start, it should be mentioned that in the previous test with the covariance function 3.2 we saw that there is no big connection between the differences of the daily solutions. So the estimated atmosphere depends only on the 'starting day' and the used difference for that day. The tests, which are done here will show, if this is a big problem for the estimation of the atmosphere. So at first, the calculated correlation, between the new calculated atmosphere and the given atmosphere is shown below:

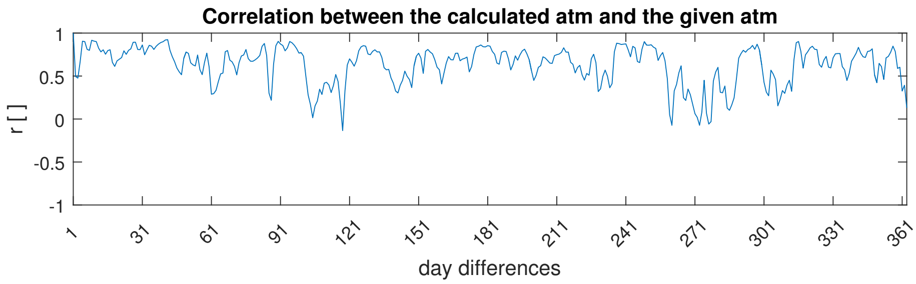



***Figure 3.20:*** *The correlation coefficients over one year of the calculated atmosphere and the given atmosphere*

A look at the numbers of the graph in figure 3.20 shows that 277 of the coefficients are higher than 0.5 and 305 are higher than 0.4, which is not a really high number, but regarding the fact that the correlation of the estimated atmosphere only depends on the correlation of each day, we can say it is a good number. It is also an improvement to the covariance test in which we tested the differences of the daily solutions and the atmosphere, in which we had only 255 coefficients higher than 0.5. Also, the mean value 0.6301 is quite good. At the moment, it can be said that regarding the bad result in the covariance function, the result has a higher correlation as perhaps expected. Of course, we can also calculate the correlation, by taking the spherical shape of the earth into account, which creates the following results:

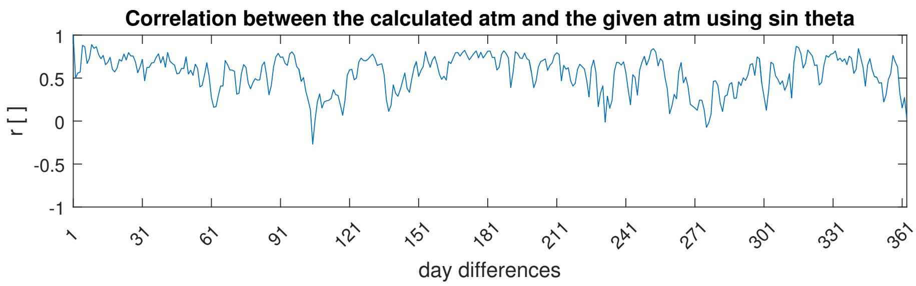
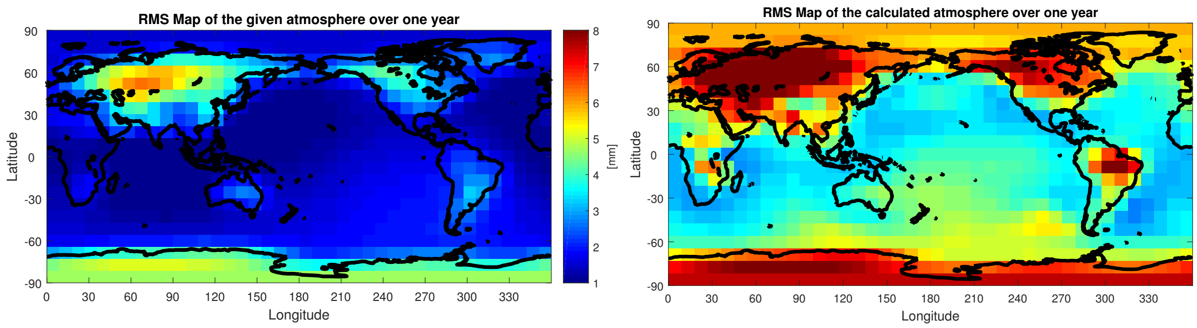



***Figure 3.21:*** *The correlation coefficients over one year of the calculated atmosphere and the given atmosphere, by regarding the sin$\theta$ values*

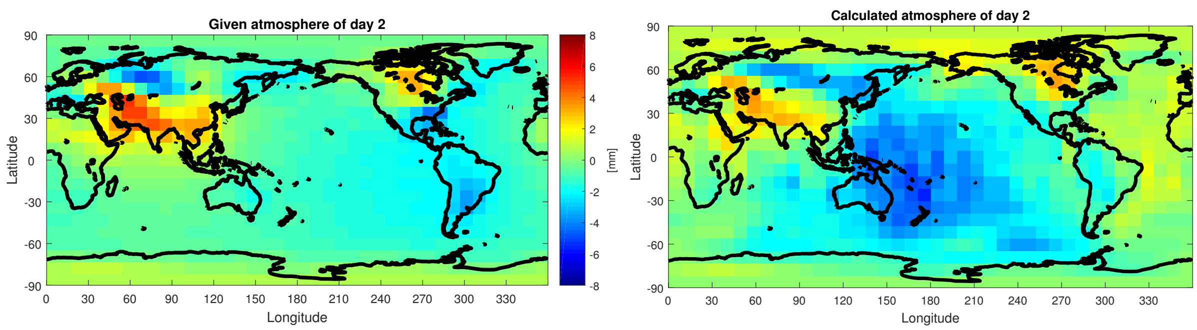
In this case we have 239 coefficients over 0.5 and 284 over 0.4, with a mean coefficient over one year with 0.5546. Also, this results show a high correlation, beside the fact that the coefficients are lower compared to the previous ones. After these two tests, it can be said that the previous method of using *masks* had a better result, but we just started to analyse this method and we can maybe improve it. In a next step, the comparison of the RMS maps will be done, to see if there are similar pattern:




(a) RMS map of the given atmosphere over one year (b) RMS map of the calculated atmosphere over one year

***Figure 3.22:*** *Comparison of the RMS maps of the given and calculated atmospheres*

In figure 3.22 it can be seen that the calculated atmosphere has much more strong positive patterns than the given one. If we compare the positive pattern over Russia and North America of the calculated atmosphere, it matches with the positive pattern in the same area of the given atmosphere, beside it is more powerful. Now we can also have a closer look at the calculated atmosphere of a single day. For this example, day 7 is chosen, the first atmospheric day, which is calculated by the differences:




(a) Given atmosphere of the 7th day (b) Calculated atmosphere of the 7th day

***Figure 3.23:*** *Comparison of the given and calculated atmosphere of the 7th day*

As it can be seen in figure 3.23, there are several similarities between the fields. For example the positive pattern over the Arab peninsula and the Middle East match match with each other. Also the positive pattern over Canada and the negative pattern over Russia can be seen in both fields. But on the other side, the strong negative pattern over Australia and the Pacific can not be seen in the given atmosphere as well as the weak positive pattern which can be seen in the Atlantic ocean. As a conclusion it can be said, that the striking areas from the given atmosphere can be seen in the estimated atmosphere, but on the other side there are several pattern in the calculated atmosphere, which can not be seen in the given atmosphere.

#### 3.4.2.2 Correlation with epochs of the atmosphere

Another possible test could pursue the question, if the calculated atmosphere match with one daily epoch of the atmosphere better, compared to the atmosphere from the whole day. In this case, the epoch three which is from 12 o'clock to 18 o'clock, epoch four which represents the time from 18 o'clock to 24 o'clock, and a combination of both will be tested. As mentioned earlier, previous tests have shown that the first two epochs have no noteworthy correlation between them and are therefore not included in this comparison. Thus, only the last two and a combination of them are used in this comparison.

***Table 3.7:*** *Comparison of the epochs with the new atmosphere*

| | 11 3rd epoch | 4th epoch | 3rd and 4th epoch |
|---|---|---|---|
| Correlation coefficients > 0.5 [ ] | 278 | 262 | 268 |
| Correlation coefficients > 0.4 [ ] | 307 | 294 | 300 |
| Mean correlation coefficient [ ] | 0.6372 | 0.6045 | 0.6236 |

Now, the spherical shape of the earth is regarded in the calculation:

***Table 3.8:*** *Comparison of the epochs with the new atmosphere regarding the shape of the earth*

| | 3rd epoch | 4th epoch | 3rd and 4th epoch |
|---|---|---|---|
| Correlation coefficients > 0.5 [ ] | 245 | 213 | 232 |
| Correlation coefficients > 0.4 [ ] | 280 | 252 | 269 |
| Mean correlation coefficient [ ] | 0.5734 | 0.5115 | 0.5720 |

After these tests, it can be said that the third epoch has the best results with 282 coefficients higher than 0.5 and 310 higher than 0.4, which means that with this kind of calculation, it is possible to reconstruct the atmosphere from 12 to 18 o'clock in a good way. In a next step, we can think about to improve this technology, one possibility is, if we remember that during the estimation it only depend on the actual daily solution and the 'starting day'. Thus, we can increase the number of 'starting days', to test, if there is also an improvement in the estimated atmosphere.

### 3.4.3 Using two fields for the calculation

The next idea is, to use two atmospheric fields, one as a start field, on which the daily differences can be added, and another after the first half of the year, to correct errors, which occur because of two effects. Because the new calculated atmosphere depends only on the start field and on the daily difference, a difference occurs. If we choose two start fields, the second one is perhaps more correlated to the atmospheres of the second half of the year. The other reason is that many different effects are included in the daily solutions, which affect the gravity field and lead to errors in the estimated atmospheres. Therefore, in a first step, it will be tested, if there is an increase of the correlation if a field is included at the first of July.

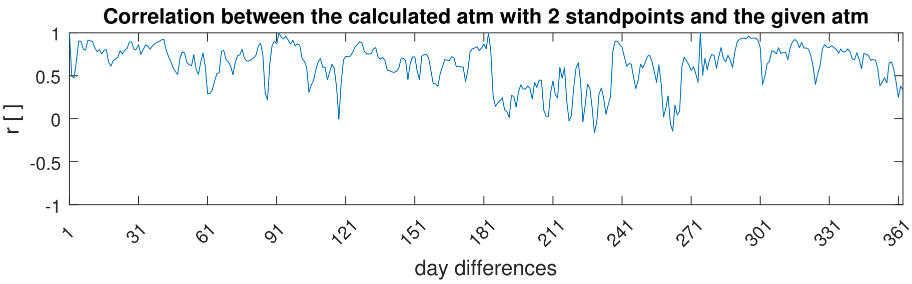
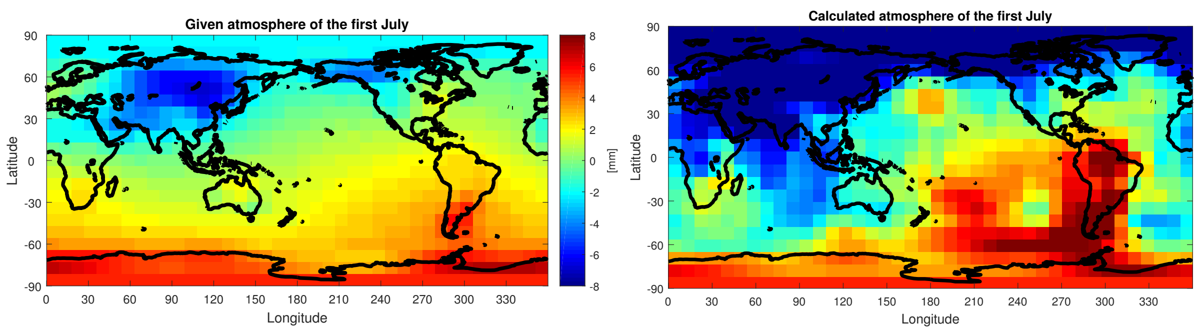



***Figure 3.24:*** *The correlation coefficients over one year of the calculated atmosphere with two fields and the given atmosphere*

A look at this graphic shows that the correlation falls to a low value, after the given atmospheric field is included at day 183, which is the first of July. After the fall to the low value, the coefficients remain in this scale area for nearly two month. That is very surprising and interesting, because according to the covariance function, there should be no big correlation between two following days. But as we can see here, it seems like they are related to each other, but sadly not in a way as we wish them to be. To see the reason for the falling of the coefficients, we can have a closer look at the calculated atmospheres. At first, we compare the given atmosphere of the first of July and the calculated atmosphere where one start field is used. This is done, because until the implementation of the second field, we have the same atmosphere as in the previous case:




(a) Given atmosphere of the first July

(b) Calculated atmosphere of the first July

***Figure 3.25:*** *Comparison of the given and calculated atmosphere of the first July*

Here we can see that the given atmosphere has a much smaller variation of the values as the calculated one. These are the two fields, where the following field will be added, to calculate the atmosphere of the following day, the second of July:

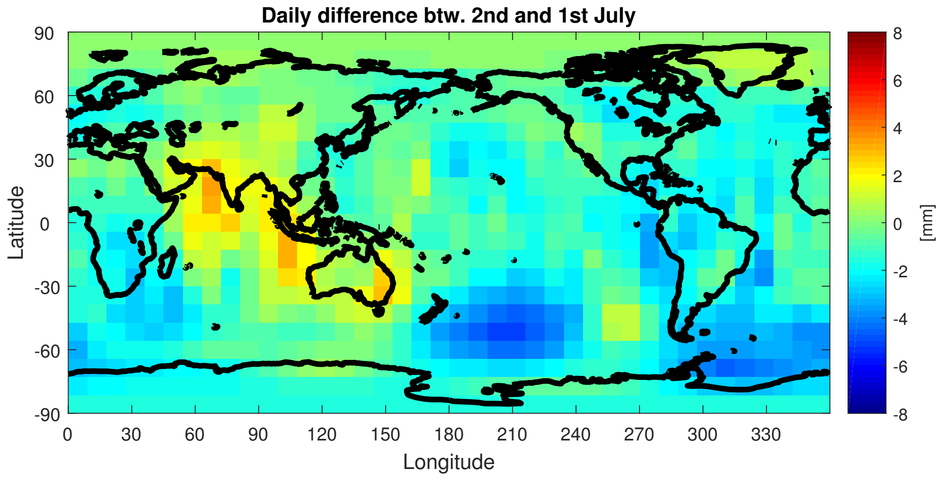
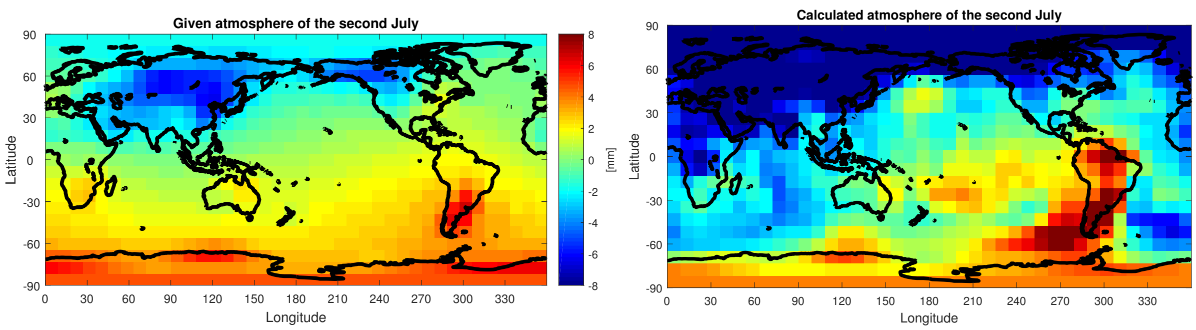



*Figure 3.26: Difference between the daily solution of the second and first July, which is used to calculate the atmosphere of the second July*

It can be seen that at the upper part a positive pattern will be added, at the lower part a mix of positive and negative areas occur. Now we can see, how this affects the atmosphere of the second July:

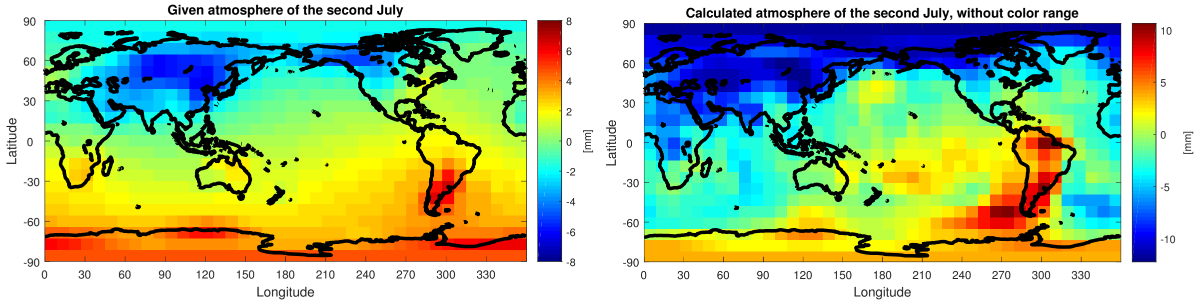



(a) Given atmosphere of the second July

(b) Calculated atmosphere of the second July, using one field

*Figure 3.27: Comparison of the given and calculated atmosphere of the second July, using one field*

Here, the atmosphere changes as expected, the upper areas show now less intensity. The problem, which occurs is that the variations of the calculated atmosphere can not be seen in the current range of the colorbar, so in this case it is helpful, to make another comparison, where the two fields, which were compared here have different colorbars to see the differences. It can already said in the figure above that the calculated atmosphere shows much stronger positive and also negative patterns than the given one.

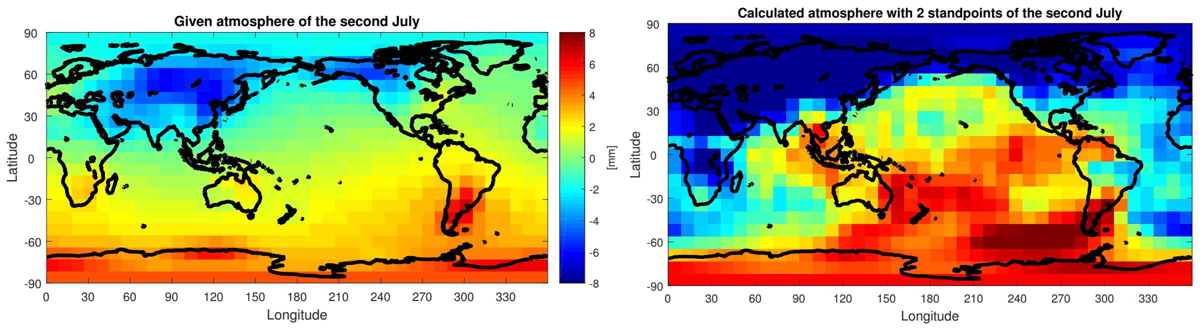



(a) Given atmosphere of the second July

(b) Calculated atmosphere of the second July, using one field, without color range

***Figure 3.28:*** *Comparison of the given and calculated atmosphere of the second July, without color range*

Again it is to mention that we have now two different colorbars. But the change has the expected positive result, so that it is now possible to see the smaller variations in the atmosphere. It can be seen that there are indeed many similar patterns between the given atmosphere and the calculated one, but if we look at the scale of the colorbar, it can be seen that they show a much larger variation as the given atmosphere. For the orientation, this comparison shows a correlation coefficient of 0.7537 and with the regard of the spherical shape of the earth a coefficient of 0.7406, which are both good results. Now, we can have a look to the comparison of the given atmosphere and the calculated one, with two fields, both from the second of July:




(a) Given atmosphere of the second July

(b) Calculated atmosphere of the second July, using two field

***Figure 3.29:*** *Comparison of the given and calculated atmosphere of the second July, using two fields*

If we remember the fact that in the differences, which were added to the given atmosphere of the first of July was a big positive pattern, we can see here the expected result that there is a bigger positive pattern. But beside that, it is still possible to identify similar patterns in both atmospheres. Again, as a comparison, for this two fields a correlation coefficient of 0.7804 is calculated and the coefficient with the regarding of the atmosphere is even bigger with 0.7944. Now, we can see in the following figure the subtracted field, where the calculated atmosphere with one field is subtracted by the calculated atmosphere with two fields:

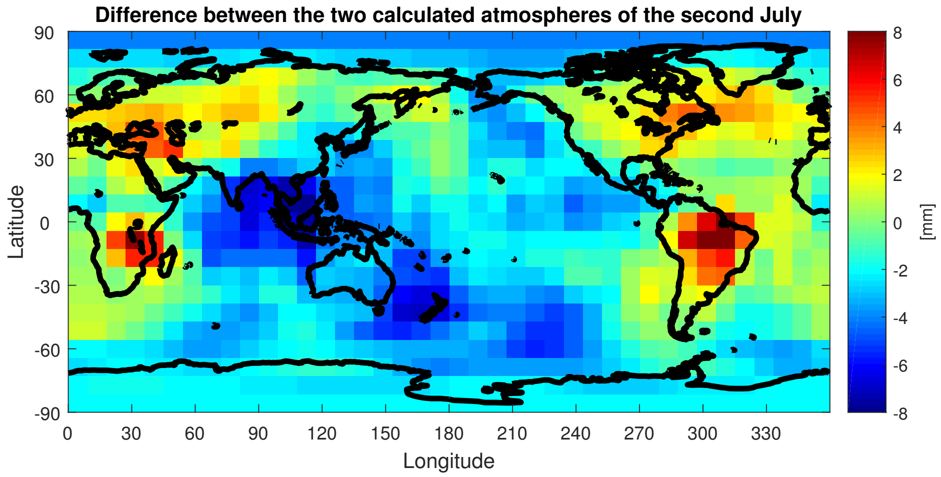



***Figure 3.30:*** *Difference between the two calculated atmospheres of the second July*

The interesting fact in figure 3.30 is that nearly no changes occurs compared to the calculated field where one atmospheric field was used. The reason is that the values in the other field are too small, to create big changes. In the next figure, the daily differences can be seen, which were used to calculate the atmosphere of the third of July:

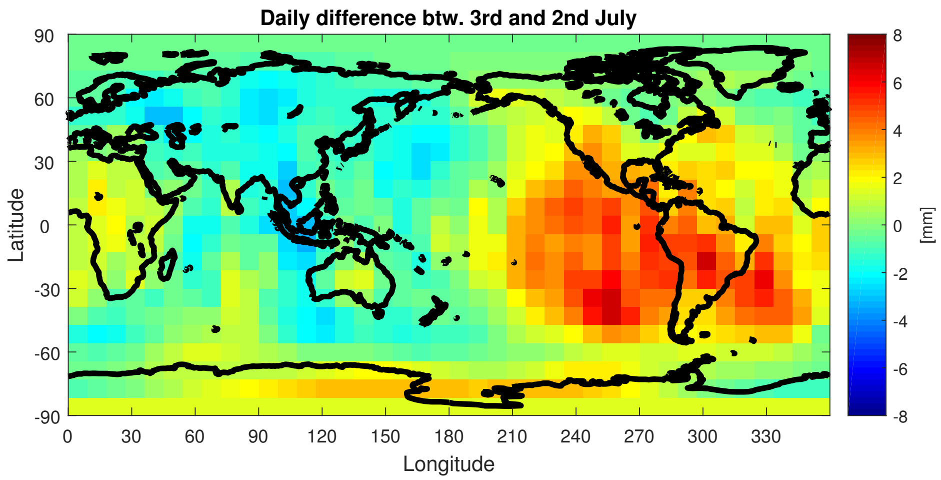
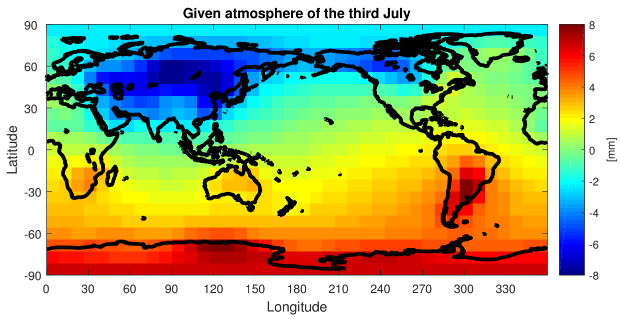
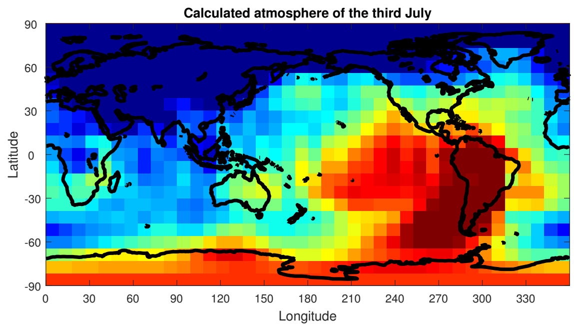



***Figure 3.31:*** *Difference between the daily solution of the third and second July, which is used to calculate the atmosphere of the third July*

Because we saw in the last figure that the colorsise is different in the two ways of estimated atmosphere, we should also have a look at the size. As it can be seen, it is more or less between 3 and -5. It also becomes apparent that now the eastern part of the atmosphere shows a large negative pattern, while only on the west side a positive pattern occurs. In the next step, we will have a look at the effect in the new calculated atmosphere:



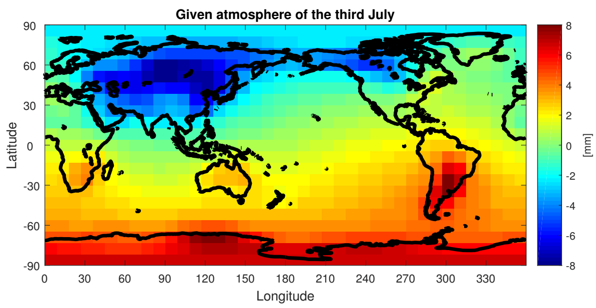

(a) Given atmosphere of the third July



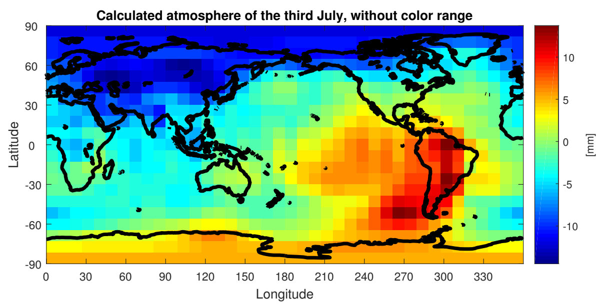

(b) Calculated atmosphere of the third July, using one field

***Figure 3.32:*** *Comparison of the given and calculated atmosphere of the third July, using one field*

Here, where only one atmosphere was used for the calculation, is it possible, to see some characteristic patterns in both fields. Again, it is helpful in this case to change the size of the correlation in the calculated field, to see the smaller variations in both images:



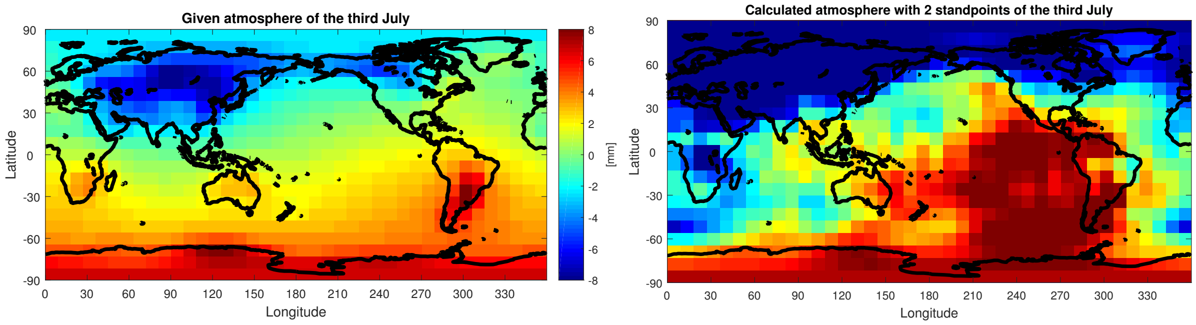

(a) Given atmosphere of the third July




(b) Calculated atmosphere of the third July, using one field, without color range

***Figure 3.33:*** *Comparison of the given and calculated atmosphere of the third July without color range*

With this change, it is now possible to see the different patterns in the calculated atmosphere better and it can be said that there are still some similar patterns, beside the fact that the size of the variation increased compared to the given atmosphere. Because of the fewer similarities compared to the previous comparison it is no surprise that the correlation coefficient in this case is with 0.6510 lower as the previous one and with 0.5932 regarding the shape of the earth also much lower. Now we can see the atmosphere which is calculated with two fields, and where the correlation coefficients indicated a big decrease of correlation to the given atmosphere:




(a) Given atmosphere of the third July

(b) Calculated atmosphere of the third July, using two field

***Figure 3.34:*** *Comparison of the given and calculated atmosphere of the third July, using two fields*

Here, in figure 3.34 it can be seen, why the correlation coefficient is much smaller than the previous ones. Because of the adding of the field of differences, the positive pattern in the right area is now divided into two fields because a negative pattern occurs between them. Also, the right one is dizzy and the left field occurs much stronger. The problem now is that in the given atmosphere it is the exact opposite. While in the right area the atmosphere has a positive pattern it is weak in the left area. So it can be said that the changes are much bigger than they should be. Here, the correlation is only 0.2682, and with the regarding of the shape of the earth it is 0.1556. Also, when we have a look a the differences between the two calculated atmospheres it can be seen that the new calculated atmosphere with two fields has less influence to the other calculated atmosphere, which can be seen in the following figure 3.35:

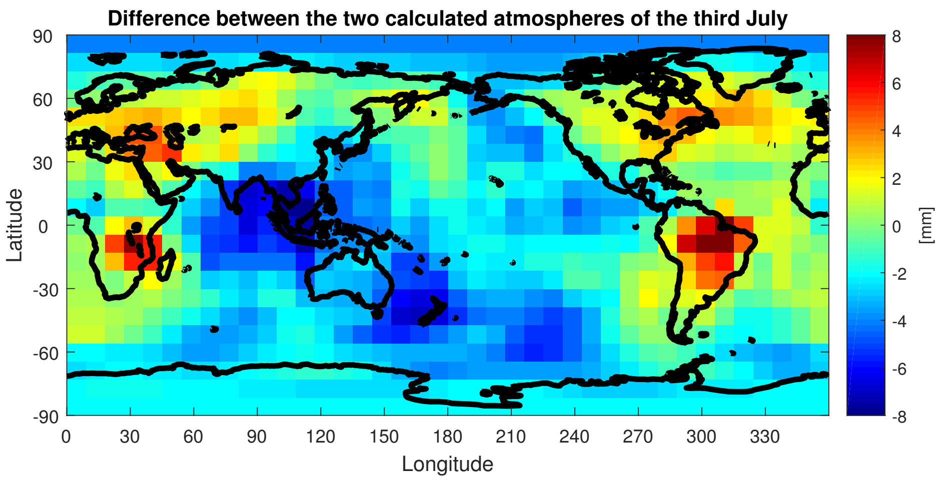



***Figure 3.35:*** *Difference between the two calculated atmospheres of the third July*

The conclusion of this comparison is that the differences of the daily solutions, which were added to create a new atmosphere, show a larger scale as the used atmospheric fields, which were given. Thus, it is clear that if the atmosphere is more weak, the strong patterns of the differences which were added have a bigger impact on the current pattern, so that they change the atmosphere more than it would change in reality that means in other words that the calculated atmosphere has less similarities with the real atmosphere of that day. The reason, why this happen in July and not at the beginning of the year, is because of the longer and warmer days, which heat up the oceans and the landmasses, according to the paper (*Atmospheric pressure, winds, and circulation patterns, p. 113 - 137*, n.d.). One solution, which is easy to implement, is that the scales of the fields of the daily differences can be reduced, for example by the half. This can be done through the summer month, where the solar heating creates a weak atmosphere. With this, the atmospheric changes can be estimated, but with the lower strength of the differences, no additional changes will be created.

#### 3.4.3.1 Estimating the atmosphere with two fields and different weighted fields of difference

With the new calculated atmosphere, we receive the following correlation, compared with the given atmosphere:

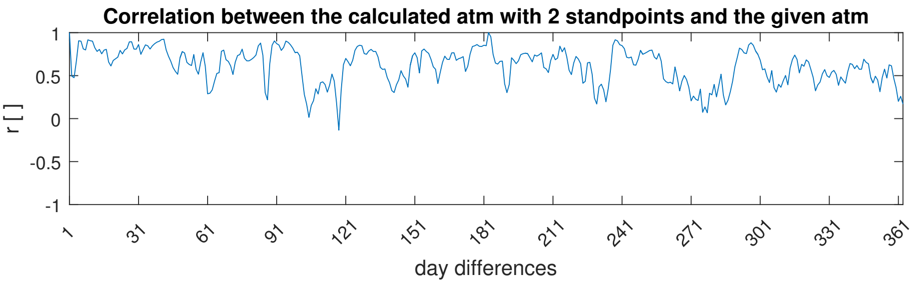



***Figure 3.36:*** *Correlation coefficients for one year, by using two fields for the calculation*

In this case, we have at 259 days coefficients higher than 0.5 and at 301 days coefficients higher than 0.4. The mean coefficient is 0.6117, which is a not very motivating result. It can be already said that in this case there is no real improvement compared to the usage of one field. We can also have a look at the calculation of the correlation coefficients, where the shape of the earth is regarded:

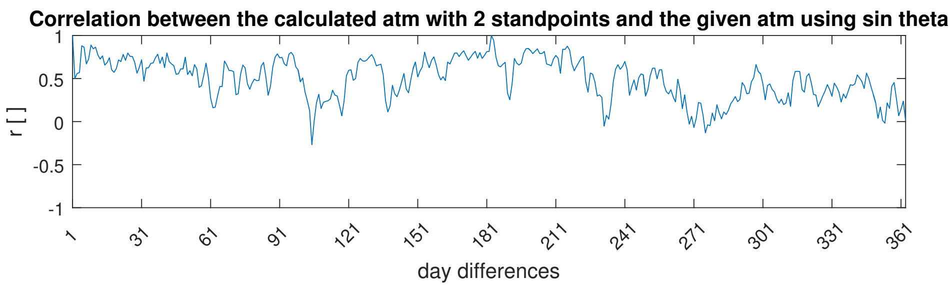
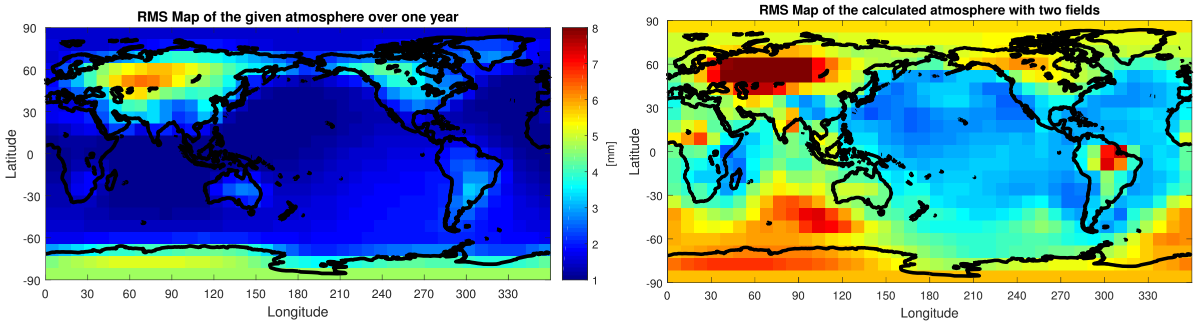



***Figure 3.37:*** *Correlation coefficients for one year, by using two fields for the calculation and regarding the shape of the earth*

Here we have at 197 days a correlation coefficient higher than 0.5 and at 241 days coefficients higher than 0.4. The mean coefficient is now 0.4977, which is a bit below 0.5, so not a really good value. Finally, we have a look at the comparison of the RMS maps of the actual estimated atmosphere and the given atmosphere:




(a) RMS map of the given atmosphere over one year

(b) RMS map of the calculated atmosphere over one year

***Figure 3.38:*** *Comparison of the RMS maps of the given and calculated atmospheres*

As we can see, the two fields show some similarities, for example over Russia and North America. But there are also many areas which show different patterns. The most significant point is that some areas are shown strong positive patterns, where in the RMS map of the given atmosphere are weak negative patterns. As a conclusion it can be said that the use of two fields lead to no real improvement of the estimation of the atmosphere. In the opposite, the results are worse than in the case when using one field. Another possibility is, to use four fields in three month steps, which will be tested in the following chapter.

### 3.4.4 Testing the estimated atmosphere where 4 fields where used

In the previous test we have seen that there is no improvement to use two atmospheres, so it could be helpful, to try to improve the previous results by using more start fields. Now we will test if there is an improvement by using four fields and, as done before, we will give the fields different weights. After some tests of combinations, the best result of weighting the differences of the daily solutions is, to use the differences of the month January to June normal, then use half of the differences during the months July to October, and the differences of the months November and December again normal. The result of this combination can be seen below:

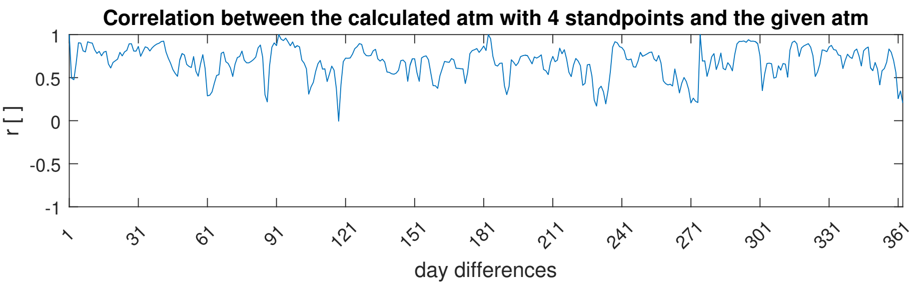



***Figure 3.39:*** *Correlation coefficients for one year, by using four fields for the calculation*

Here, 310 coefficients are higher than 0.5, which is a high number and with 335 coefficients higher than 0.4, the result is much higher. Thus, we have already an improvement to the previous cases where we only used one or two fields. The mean coefficient over one year is 0.6796.

This result is good and it is also a big improvement compared to the previous results. In the next step, we have a look at the correlation coefficients which were calculated with regard to the spherical shape of the earth:

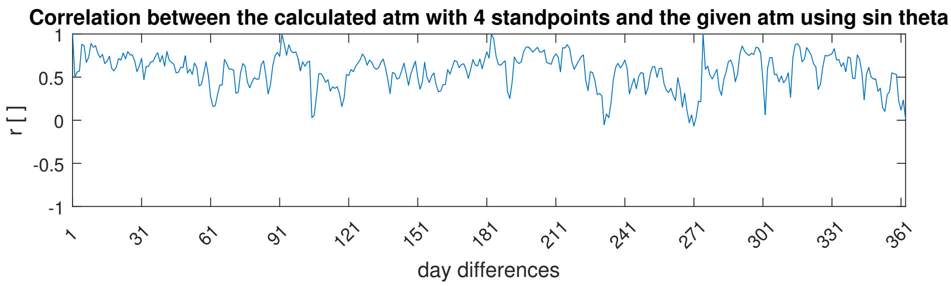
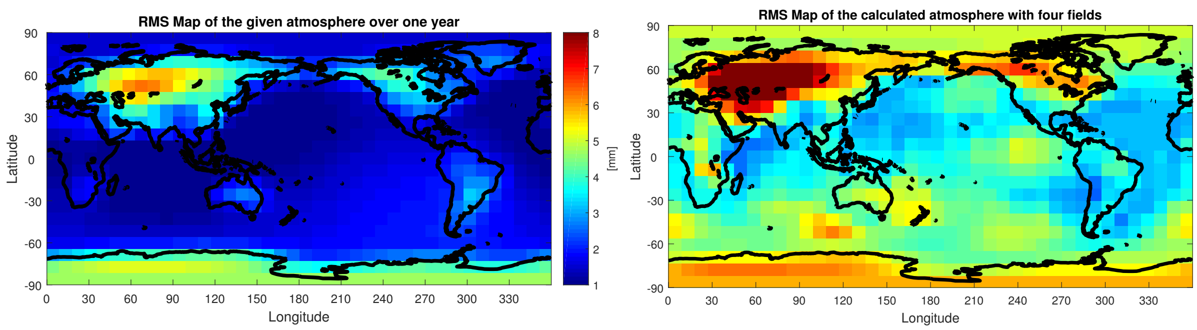



*Figure 3.40:* *Correlation coefficients for one year, by using four fields for the calculation, regarding the shape of the earth*

In this case, 250 coefficients are higher than 0.5 and 295 are higher than 0.4, which are again good results. The mean coefficient is with 0.5699 also high, so it can be said that also these numbers show a high improvement. Finally, we compare the RMS map over one year from the given atmosphere with the estimated atmosphere:




(a) RMS map of the given atmosphere over one year

(b) RMS map of the calculated atmosphere over one year

*Figure 3.41:* *RMS maps over one year of the given atmosphere and the calculated atmosphere*

If we have a look at these maps in figure 3.41 we can see that there are several similarities, but on the other hand there are still areas, which have different patterns. Compared to the usage of only two fields, it can be said that here is again an improvement. Now we can at least use atmospheric fields on each month, to improve the results:

### 3.4.5 Estimating of the atmosphere with 12 fields

In a first step, we can test, if there is an increase, if we use instead of two fields, 12 atmospheric fields, each at the beginning of every month.

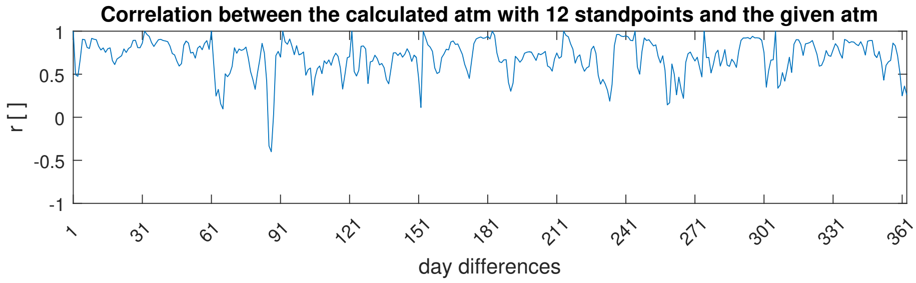



***Figure 3.42:*** *The correlation coefficients over one year of the calculated atmosphere with twelve fields and the given atmosphere*

Here, 313 coefficients are higher than 0.5 and thus, higher than the previous one and 332 higher than 0.4, which is a bit lower than in the previous test where 335 coefficients were higher than 0.4. The mean coefficient is 0.6989, so it is higher than the previous one with 0.6796. After this test it can already be said that this method is also useful to estimate the atmosphere. If we have a look at the correlation, where the shape of the earth is regarded, we have this result:

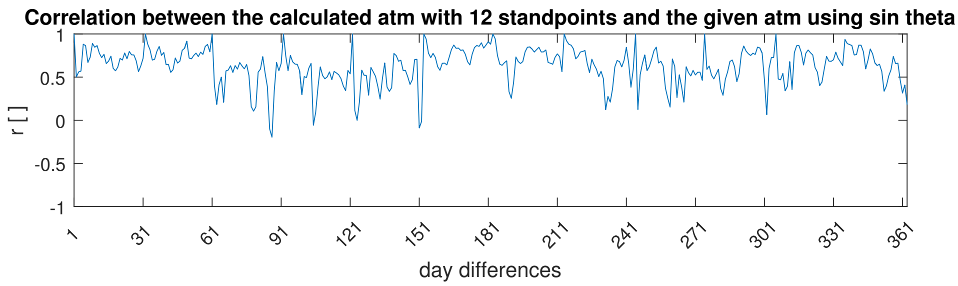
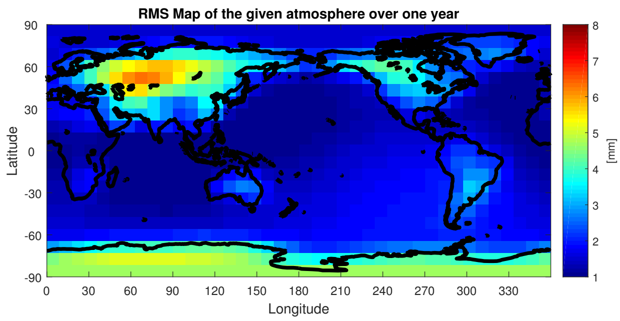
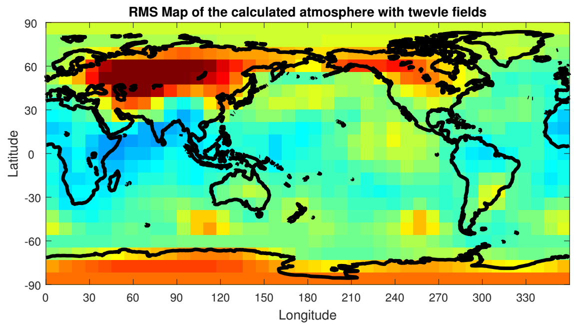



***Figure 3.43:*** *The correlation coefficients over one year of the calculated atmosphere with twelve fields and the given atmosphere, regarding the spherical shape of the earth*

In this comparison in fugure 3.43, 290 coefficients are higher than 0.5 and 313 coefficients are higher than 0.4 which are higher than the previous tests. At last, we can have a look at the comparison of the RMS map over one year from the given atmosphere and the RMS map of the actual calculated atmosphere:




(a) RMS map of the given atmosphere over one year




(b) RMS map of the calculated atmosphere,by using the differences between the daily solutions and using twelve atmospheric fields

***Figure 3.44:*** *Comparison between the RMS maps of the calculated atmosphere by the daily differences and the use of twelve atmospheric fields and the given atmosphere*

A look at figure 3.44 shows, that there are many similarities and it can be said, that this is the best result compared to the other estimated atmospheres by using start fields. In some areas are still different pattern, but they are small compared to the previous results. After this test, we will have a look at further possibilities to analyse the estimated atmospheres.

### 3.4.6 Further analyses of the atmospheres

#### 3.4.6.1 Power spectrum of the calculated atmospheres

Until now, we only had a look at the correlation coefficients and the RMS maps to analyse the calculated atmospheres. To test the impression of the different power of the coefficients, it is also possible to show graphics of the degree RMS, which show the RMS over the degree, as done in previous tests by using the formula ( 2.26). In the following example, the degree RMS of the 300th day of the atmosphere, calculated with one, two, four and 12 start fields is shown. We will choose day 300 to see the differences, if we regard the fact that some atmospheres start every month or after half a year, with this day we will have a good overview, how the atmospheres look after some distance to the start field:

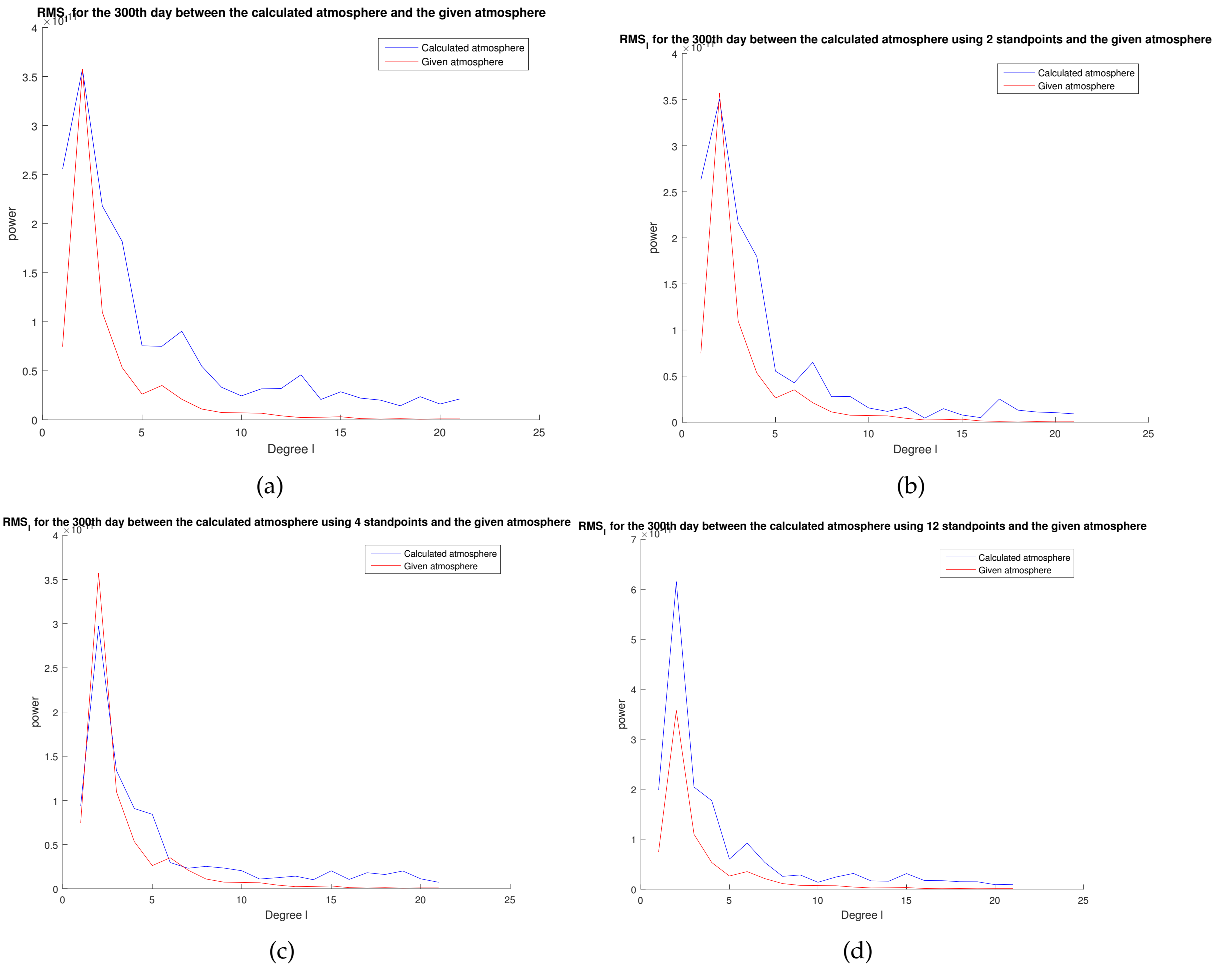
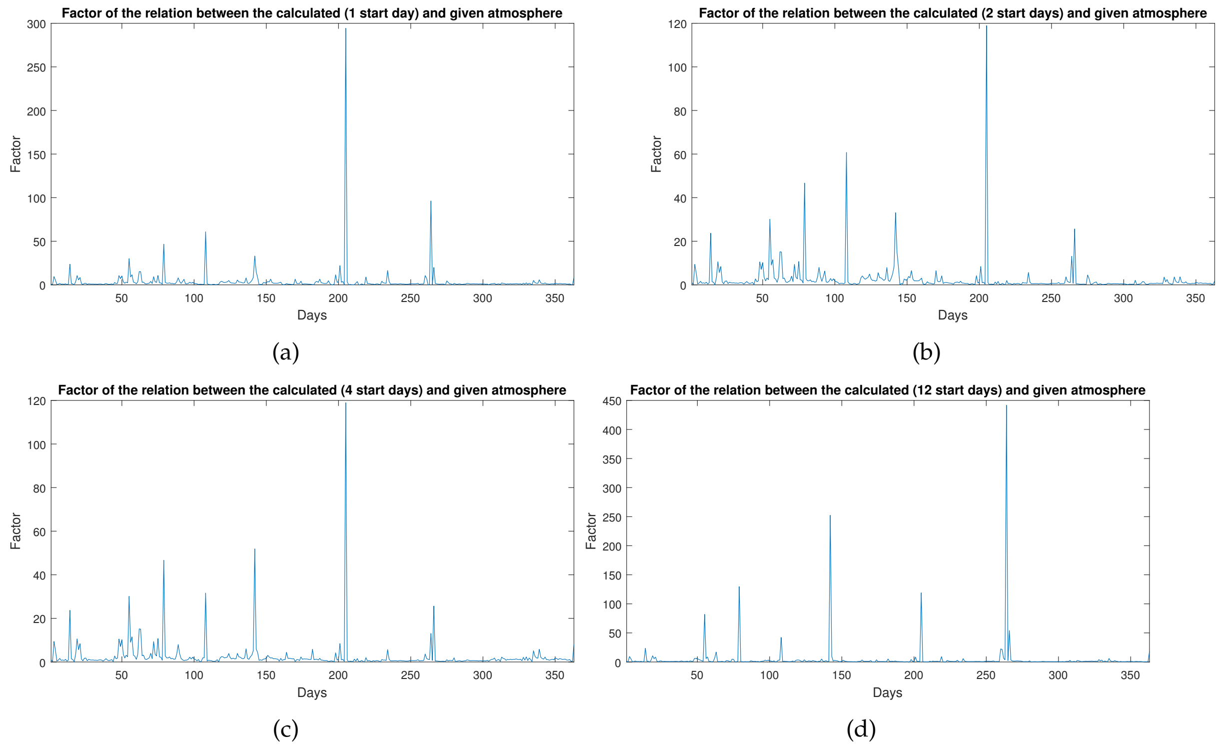



***Figure 3.45:*** *Power spectra of the estimated atmospheres*

The first point which stands out by looking at the graphs in figure 3.45 is that in opposite to the previous tests, where we had a stronger peak from the graph of the estimated atmosphere, here in three of four cases, the given atmosphere has the same peak or is even stronger. Interestingly, the given atmosphere is stronger in comparison with the estimated atmosphere by using 4 start fields. In the cases, where one or two start fields are used, the calculated and given atmosphere has the same strength, which are both the comparisons, where the calculated atmosphere has the longest distance to the start field. The atmosphere which was calculated with start fields every three month shows a lower peak compared to the given atmosphere for the lower coefficients. Finally, the atmosphere with the monthly start fields shows a stronger peak compared to the given atmosphere. Because of the fact that in the low degree coefficients the atmosphere is strong, we know that the low frequencies of the atmosphere have a high energy, which decreases faster compared to the energy of the calculated atmospheres, which have at a higher frequency and degree more energy than the given atmosphere. Therefore, this could be one source of errors, as already explained in the previous comparison, it would be very interesting to make further analyses of this effect in further studies.

#### 3.4.6.2 Testing the atmosphere on a factor difference

As we saw in the estimation of the atmospheres by using *masks* (figure 3.19) it is also possible that there is a factor difference, which leads to a high peak in the intensity of the estimated atmospheres. Now the interesting question is, if in this kind of calculation also such peaks appear. To answer this question, the formula ( 3.10) will be used for the calculation of the factors:




(a) (b) (c) (d)

***Figure 3.46:*** *Graphs of the factors of the estimated atmospheres*

If we have a look at the graphs in figure 3.46, two things can be seen. At first, it becomes obvious that, as expected, the high peaks appear at the same days as in the previous case. Again, it is to mention that this effect is worth a further analyse, which would extend this work. But the main question is, why the peaks appear at these special days. One point, which is very interesting, that we have to keep in mind that we reduced the strength of some of the fields about the half for example. That was done, beside others, at the atmosphere where we used four start fields. We reduced the intensity of the differences of the atmosphere about the half, for the month July to December. But the peak at the 23th July is very prominent. It has the same strength compared to the peak by the atmosphere, where we used the *masks*. This is also a point which needs to be analysed closer, because it could be expected that the peak should be also reduced. The second effect, which can be seen is the increase of the peaks' strength when we add more start fields. This is also an effect, which is difficult to explain without further tests. If we regard the fact that each new calculated atmosphere depends mostly on the start field and the used difference, the increase of the peaks would mean that a closer atmospheric day to the month in which the peak appears leads to more energy. This is a very interesting effect, but a close analyse would extend the size of this thesis, further analysis can be done here, too. The following table shows the number of the factors below 2, for the different atmospheres, to get an idea of the strength of the peaks:

***Table 3.9:*** *Number of factors below 2*

| | Atmosphere using 1 start field | Atmosphere using 2 start fields | Atmosphere using 4 start fields | Atmosphere using 12 start fields |
|---|---|---|---|---|
| Factors below 2 | 250 | 271 | 287 | 293 |

This table 3.9 leads to the assumption that there are small variation which can be reduced, if we have a good estimation of the atmosphere. The reason is that the number of factors below 2 increase, when we add more start fields, thus, the small peaks which can be seen in the graphs in figure 3.46 decrease. But, in opposite to the small variations the prominent peaks increase by adding more start fields, which should also be analysed in further tests, as mentioned before.

### 3.4.7 Conclusion of the tests

As a conclusion, it can be said that the use of atmospheric fields for each month offers the best results, but if there are not so many fields available, the use of fewer fields still gives us also good results. Compared to the previous method, where we used the *masks* to estimate the atmosphere, the results are worse, but still usable. Now, it can be tested, in which way these three possible atmosphere estimations (two fields, four fields and 12 fields) match with the different epochs from the atmosphere.

### 3.4.8 Calculating the correlation for the epochs of the atmosphere

In order to offer a better visibility of the differences, only tables will be used in this part. In a first step, the third epoch from 12 - 18 o'clock will be analysed:

***Table 3.10:*** *Number of coefficients for epoch 3*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 263 | 313 | 317 |
| Correlation coefficients $> 0.4$ [ ] | 298 | 335 | 335 |
| Mean correlation coefficient [ ] | 0.6195 | 0.6869 | 0.7053 |

If the shape of the earth is regarded during the calculation, the following results for the third epoch can be seen:

*Table 3.11: Number of coefficients for epoch 3 regarding the shape of the earth*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 206 | 255 | 292 |
| Correlation coefficients $> 0.4$ [ ] | 244 | 298 | 321 |
| Mean correlation coefficient [ ] | 0.5092 | 0.5815 | 0.6342 |

In the next step the correlation coefficients of the fourth epoch, which is from 18 - 24 o'clock, and the estimated atmosphere will be calculated:

*Table 3.12: Number of coefficients for epoch 4*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 242 | 310 | 309 |
| Correlation coefficients $> 0.4$ [ ] | 285 | 329 | 330 |
| Mean correlation coefficient [ ] | 0.5878 | 0.6728 | 0.6903 |

In the table below, we can see the data after we calculated the coefficients regarding the shape of the earth for the fourth epoch:

*Table 3.13: Number of coefficients for epoch 4 regarding the shape of the earth*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients $> 0.5$ [ ] | 171 | 241 | 269 |
| Correlation coefficients $> 0.4$ [ ] | 218 | 288 | 304 |
| Mean correlation coefficient [ ] | 0.4507 | 0.5516 | 0.6080 |

At last the mean field from these two epochs will be used for the comparison, thus, the estimated atmospheres are compared with the given atmosphere from 12 - 24 o'clock:

***Table 3.14:*** *Number of coefficients for epochs 3 and 4*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients > 0.5 [ ] | 252 | 312 | 309 |
| Correlation coefficients > 0.4 [ ] | 299 | 332 | 330 |
| Mean correlation coefficient [ ] | 0.6064 | 0.6831 | 0.6903 |

The last comparison shows the coefficients, with regard to the shape of the earth:

***Table 3.15:*** *Number of coefficients for epochs 3 and 4 regarding the shape of the earth*

| | Estimated atmosphere with 2 field | Estimated atmosphere with 4 field | Estimated atmosphere with 12 field |
|---|---|---|---|
| Correlation coefficients > 0.5 [ ] | 194 | 255 | 312 |
| Correlation coefficients > 0.4 [ ] | 231 | 296 | 331 |
| Mean correlation coefficient [ ] | 0.4837 | 0.5712 | 0.7013 |

As a conclusion it clearly can be said that the highest correlation occurs by using the third epoch of the atmosphere where 335 coefficients are higher than 0.4. This is the same epoch, which had shown us, also during the comparison with the estimated atmosphere with one start field, the highest correlation. In this case, the use of twelve atmospheric fields offers the highest correlation. In the next chapter, we will have a short look into the frequency domain.

## 3.5 Short look into the frequency domain

In this section, the data will be analysed in the frequency domain. Beside the fact that there are a number of tests, which could be done in the frequency domain, we will focus here only on the comparison of the different frequencies because more tests would extend the size of this work. This is the first basic step for further tests and gives us already an idea, how the results can be improved or where further analysis are necessary. Before the analysis can start, the theory behind this should be explained, how it is possible to transfer a signal from the space domain into the frequency domain. This can be done by using the so called Fourier transformation, or in this case, a discrete Fourier transformation. The formula for the discrete Fourier transformation from the vector $X$ in the space domain into $Y$ in the frequency domain can be seen below (e.g. (Mat, n.d.*a*)):

$$Y(k) = \sum_{j=1}^{n} X(j) W_n^{(j-1)(k-1)} \tag{3.11}$$

Where $W_n$ is:

$$W_n = e^{(-2\pi i)/n} \tag{3.12}$$

Where $n$ is the length of the vector which will be transformed into the frequency space and $W_n$ is 'one root of unity' (Mat, n.d.*a*). To be more precise, the calculation in MATLAB was done by a fast Fourier transformation, which means that the discrete Fourier transformation is implemented in very efficient algorithms[1]. Here, it is to mention that not the whole field is transformed into the frequency domain, the transformation is done for one point through the whole year. The reason is that we only can use one vector for the transformation. Because of this, it is better to test several points around the whole field, in order to have an impression about the comparison of the daily solution and the atmosphere, in this case we have chosen four points. It has to be mentioned that for this test the daily solutions are not subtracted by the eleven day mean fields, because this would lead to an eleven day periodicity and therefore it would create wrong frequencies in the spectrum. The reason, why only the daily solutions and not the estimated atmospheres can be tested, is due to already explained reasons. In the first method to estimate atmospheres, we use *masks*, which also create a periodicity for eleven days or monthly, which can be seen in the frequency domain and would change the real signal. The other method, by using start points include the differences of the daily solutions, which also change the real signal. Therefore, in this test only the daily solutions will be tested. In the following RMS map of the atmosphere over a whole year, it can be seen which points were chosen:

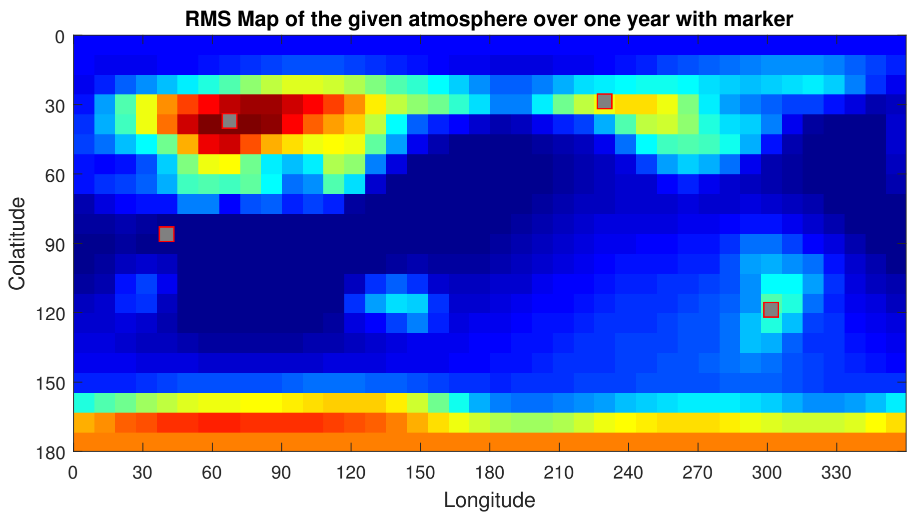
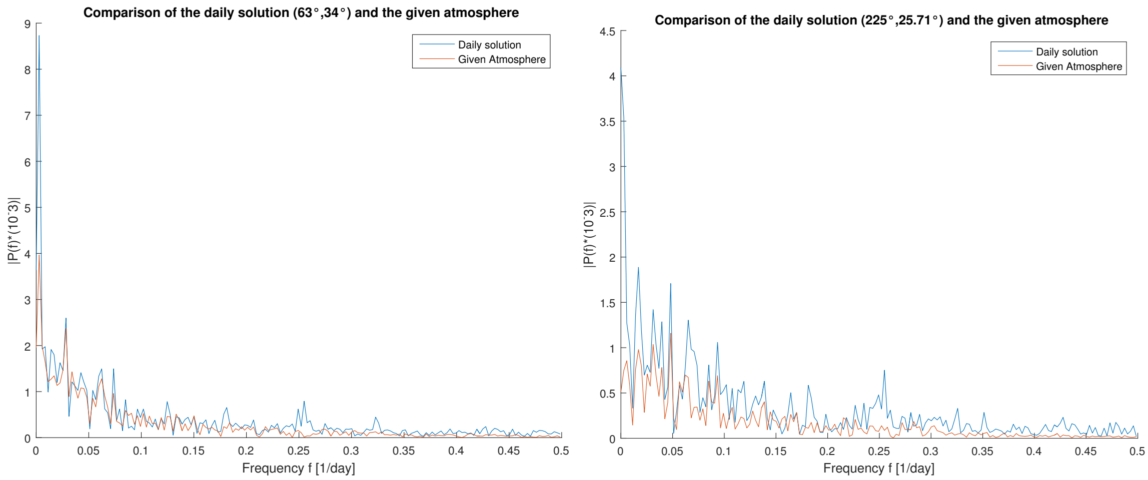



***Figure 3.47:*** *RMS map of the given atmosphere with the marker at the points, which are tested for the frequency domain*

[1](Mat, n.d.*b*)

As it can be seen in figure 3.47, the points are chosen in interesting areas, one in the most positive pattern and one in a weak positive area. To see the negative pattern, also one point in the strongest negative area is chosen and one in a very weak area, where nearly no variations occur during the year. With these points, it is possible to have a good impression about the effects of the different areas in the frequency domain. At first, we will have a look at the two points, which are located in the positive areas:

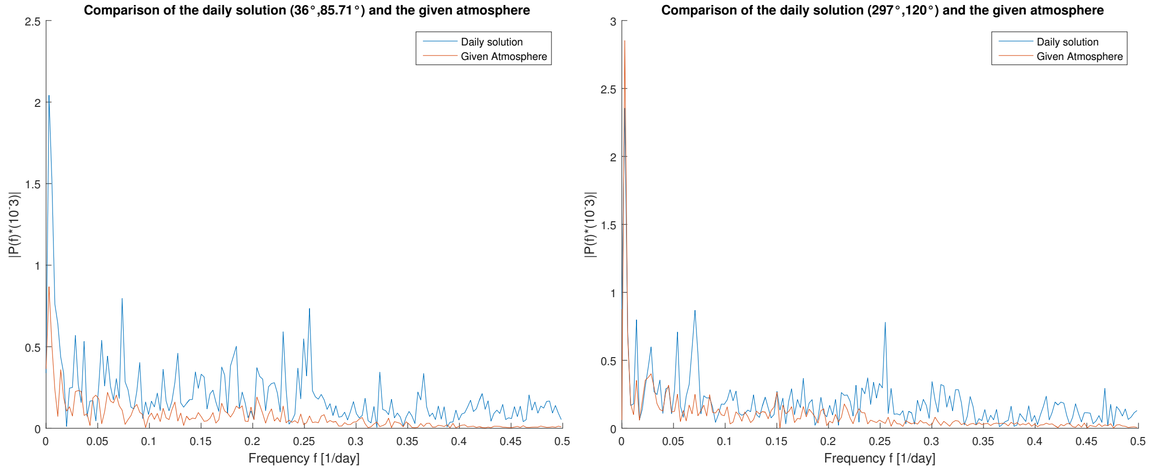



(a) Comparison of the atmosphere and the daily solution in the point at 63° longitude and 34° latitude

(b) Comparison of the atmosphere and the daily solution in the point at 225° longitude and 25.71° latitude

***Figure 3.48:*** *Comparison of the atmosphere and the daily solution in the frequency space with point in positive points*

In the left graphic in figure 3.48 it can be seen that the signals from the daily solution are stronger compared to the signals from the atmosphere, but there are several similarities between the two signals. These similar areas occur mostly in the lower frequencies. A look at the right image shows that there are not many similarities. However, we can see that in this graph the signal from the daily solution is stronger compared to the signal from the atmosphere. These stronger signals could lead to a high influence in the higher frequencies, because there the atmosphere shows nearly no signals. The reason for the stronger signals in the higher frequency space could be caused by other effects, which are comprised in the daily solutions. So, a higher correlation between the atmosphere and the daily solutions could be received, when these disruptive signals are reduced. In the next step, the two lower points will be analysed:




(a) Comparison of the atmosphere and the daily solution in the point at 36° longitude and 85.71° latitude

(b) Comparison of the atmosphere and the daily solution in the point at 297° longitude and 129° latitude

***Figure 3.49:*** *Comparison of the atmosphere and the daily solution in the frequency space with point in negative points*

A look at the two graphics in figure 3.49 shows that here the daily solution is much stronger compared to the atmosphere. Because of that, only in some areas similarities can be seen. This appears mostly in the left image for the point 36° longitude and 85.71° latitude, but also at the second point in a slight way, which leads to the result that here are more similarities. These two comparisons show that it is very helpful for further comparisons to reduce the intensity of the daily solutions mostly in the higher frequencies. With this, signals from other effects could be reduced and the correlation can be improved.

These four comparisons show that even in the unchanged daily solutions some similarities between them and the atmosphere can be seen. But they appear mostly at the sides of the peaks, the intensity of the peaks vary between them, thus it would be useful to reduce the intensity of the daily solutions, so that they match better. It can be assumed that there are similar signals but because of the fact, that in the daily solutions are also other effects included, the similar signals are partly covered by them. Thus, it should be possible to extract these similar signals to estimate the atmosphere. However, another point is that the atmosphere is very weak in the high frequency area on the contrary to the daily solutions, which have much energy in this frequency area. It could be possible to use low pass filters to eliminate the disruptive effects. In the last step, we will now have a look at the significance to test, how significant the correlation between the estimated atmospheres and the given atmosphere is.

## 3.6 Significance Tests

We have found two ways to estimate the atmosphere by using the daily solutions and the given atmosphere and the correlation is in both cases good. But the question occurs, if it is also a significant good result? If we have a short excursus to the statistic, we find the fact that not every result, which looks good, is also significant correct. To test this, we can make so called significance tests, which gives us the information to which percentage our result matches significant with the given atmosphere. In this case, the so called Wilcoxon rank sum test will be done, to analyse, if the estimated atmosphere and the given atmosphere are 'samples from continuous distributions with equal medians'(Mat, n.d.*c*). However, here we have the problem that the test requires independent test values and also values which have the same standard deviation. This creates some problems which we have to solve before the real significance test can be done. It can be imagined that the weather of one day is related to the weather of the previous day. So if a point is analysed during a certain time period, the values of the vector are related to each other. Thus, we first have to eliminate the year frequency (which is the lowest frequency in the signal) and the constant signal with the frequency 0. After this is done, only every tenth day is chosen for the vectors to eliminate also the main correlation between following days. Because it can be assumed that after ten days the fields have no high correlation which can be seen in the covariance function which was calculated earlier and can be seen in figure 3.5. Thus, we still have a low correlation between the samples but with these methods, it can be said that we reduced it to a low value. The latter point is even more difficult, here we can not make changes in our data, so we have to test, if the estimated atmosphere and the given atmosphere have the same standard deviation. This can be done by using the MATLAB function 'vartest2', which tests two samples for equal variances by using a two-sample F-test[2]. In this test, the null hypothesis suggest 'that the data in vectors x and y comes from normal distributions with the same variance'[(Mat, n.d.*d*)] and the alternative hypothesis is the opposite, thus that the two samples come from normal distributions, with different variances. By using the standard deviations $s_1$ and $s_2$ from the samples, we can write the test statistic as followed[e.g. (Mat, n.d.*d*)]:

$$F = \frac{s_1^2}{s_2^2} \tag{3.13}$$

So we use the variances of the two samples in equation ( 3.13) to calculate the ratio between them, which should be close to 1 in the case we want to accept the null hypothesis. Now we can test the daily solutions and the estimated atmospheres, if they have the same standard deviation as the atmosphere, where we will calculate the test for each of the four points, where we also calculated the frequency domain. In each test, we receive the p - value which is between zero and one and gives with a 5% significance level the information, if the null hypothesis is accepted or not. The other information which is provided as a result is the h value, which is 0 in the case the null hypothesis is accepted and 1, if it is not accepted. To make it more obvious, the p-value will be shown in the table with the h-value in bracket behind it:

[2](Mat, n.d.*d*)

*Table 3.16: Results of the vartest of the estimated and the given atmosphere*

| Point coordinates | Daily solution | Atm. using 11 day *masks* p(h) | Atm. using monthly *masks* p(h) | Atm. using 1 field p(h) | Atm. using 2 fields p(h) | Atm. using 4 fields p(h) | Atm. using 12 fields p(h) |
|---|---|---|---|---|---|---|---|
| (63°, 36°) | 0.1401 (0) | 0.5775 (0) | 0.0399 (1) | 0.1236 (0) | 0.2562 (0) | 0.0713 (0) | 0.1117 (0) |
| (225°, 25.71°) | $6.0730{\cdot}10^{-6}$ (1) | $2.7085{\cdot}10^{-6}$ (1) | 0.7933 (0) | 0.0199 (0) | 0.0198 (1) | 0.0088 (1) | 0.0128 (1) |
| (36°, 85.71°) | $5.1976{\cdot}10^{-7}$ (1) | $4.8095{\cdot}10^{-20}$ (1) | $3.4566{\cdot}10^{-6}$ (1) | $1.0648{\cdot}10^{-5}$ (1) | $7.1887{\cdot}10^{-7}$ (1) | $1.3955{\cdot}10^{-8}$ (1) | $6.8957{\cdot}10^{-7}$ (1) |
| (297°, 120°) | $2.2174{\cdot}10^{-4}$ (1) | $1.4544{\cdot}10^{-15}$ (1) | 0.0088 (1) | $5.6329{\cdot}10^{-5}$ (1) | $5.5305{\cdot}10^{-4}$ (1) | $6.6409{\cdot}10^{-4}$ (1) | $1.7344{\cdot}10^{-4}$ (1) |

A look at table 3.16 shows a very interesting result, because the only point, in which the estimated atmosphere (with the daily solution) and the atmosphere have the same standard deviation is the first point (63°, 36°). If we start with the test of the daily solution and also with the two estimated atmospheres by using *masks*, it can be seen that the first point can be used for further tests with the estimated atmosphere using 11 day *masks* and the second point can be used for the estimated atmosphere using monthly *masks*. The lower two points have no similar standard deviation. The fact that the lower points of the estimated atmospheres have no acceptance of the null hypothesis is clear, because they even do not have the acceptance of the null hypothesis in the daily solution, which is the basis of the estimation. The interesting point is that the p-value decrease very fast between the first and last two points. Because of that, it is clear that the lower two points can not be used in the significance test. At least, we can have a look again at the frequency domain, to see if there are some hints, why the results are so bad in the lower two points. So a look at the figure 3.49, gives indeed a hint, why there could be another deviation, if we compare these frequency domains with the one of the two upper points in figure 3.48. In the upper points the daily solution frequencies and also the frequencies from the atmosphere show many similarities in their shape, however there are some peaks which do not match. In the opposite to the upper points, the lower two points show fewer similarities as already explained. So it can be assumed that here are some reasons for the mismatching, because the standard deviation is not comparable. Here should be done further analyses, why there is a big difference. The estimated atmospheres, where the atmospheric start field was used to add the differences of the daily solutions, show a worse result: Here, only for the first point the null hypothesis is accepted in all four cases and the second point only in the first method, where we use one start field. But the p values are very low, which lead to the assumption that for these tests, the significance test is not recommended, beside the null hypothesis is accepted. Thus, with this test we learned that there is also a big variation in the standard deviation inside the fields which should be analysed closer. Now we can proceed with the analyse by regarding the points, which fulfil the needed requirements for the significance test.

In this step, we will now analyse the first two points from the daily solutions as well as from the estimated atmospheres with the *masks*. This test is a so called nonparametric test, which is used when two populations are tested and the used samples not depend on each other [3]. The test has the hypothesis[4]:

$$H_0 : a = 0$$
$$H_1 : a \neq 0 \quad (3.14)$$

Where the $H_0$ hypothesis says that the two samples are identical, while the $H_1$ hypothesis says the opposite. Before the Wilcoxon rank sum statistic is explained, it is to mention that this test is equivalent with the so called Mann-Whitney U-test, which is also a nonparametric test which analyse two identical samples ($X$ and $Y$), if their medians of populations are equal. In the statistic of the Mann-Whitney U-test, $U$ gives an information of the number how often a $y$ precedes in an organized arrangement an $x$, in the mentioned two separate samples $X$ and $Y$. To go back to the Wilcoxon rank sum test, the relation between these two tests is the following, where we have to regard that $n_X$ is the size of the sample $x$[5]:

$$U = W - \frac{m_X(n_X + 1)}{2} \quad (3.15)$$

Because of the used MATLAB function 'ranksum' it is only possible to compare vectors, which is the reason, why we decided to compare only the points, which we already saw in the analysis in the frequency space of the potential function. There, we chose the same point for the given atmosphere and the field we want to compare it with. For each test we receive, as in the previous test, a p - value and a h value, which is 0 in the case the null hypothesis is accepted and 1, if it is not accepted. Now, in a first test, we want to test, if there is already a significant correlation between the daily solutions (reduced by the eleven day mean field) and the atmosphere. The reason is that this is the basis of all calculations and in the frequency domain the daily solutions (in that case with the eleven day mean field included) showed good results in the comparison to the given atmosphere. Therefore, the results of the significance test can be seen in the following table:

***Table 3.17:*** *Results of the significance test of the daily solution and the given atmosphere*

| point coordinates | p - value | h - value |
|---|---|---|
| $(63°, 36°)$ | 0.9158 | 0 |

The result, which can be seen in the table 3.17 is that the null hypothesis is accepted, so it can be assumed that the two elements are from distributions with the same median. A look at the p - value show that it is very high with 0.9158, so that the null hypothesis is accepted clearly. This result is motivating because it can be expected that even without the changes in the frequency domain, the daily solutions match with the given atmosphere. If we have a look at the figure 3.48, it can be seen that the point with the high p-value (63°, 36°) also show similarities in the frequency domain. This is a good example, how the strong or weak signals

[3](Mat, n.d.*c*)
[4](*Wilcoxon-Mann-Whitney-Test*, n.d.)
[5](Mat, n.d.*c*)

in the frequency domain affect the comparison with the atmosphere. In further tests it can be helpful to reduce the energy of the signals from the daily solutions so that it match better with the signals from the correlated atmosphere. Now, in a next step, the estimated atmospheres and the given atmosphere in the point (63°, 36°) can be compared by using the significance test. To make it more clear, the p - value of each test is shown and in bracket the h value:

***Table 3.18:*** *Results of the significance test of the estimated and the given atmosphere*

| point coordinates | Atm. using 11 day *masks* p(h) | Atm. using 1 field p(h) | Atm. using 2 fields p(h) | Atm. using 4 fields p(h) | Atm. using 12 fields p(h) |
|---|---|---|---|---|---|
| (63°, 36°) | 0.7690 (0) | 0.9148 (0) | 0.7397 (0) | 0.8880 (0) | 1.0000 (0) |

It can be said that all tests have very good results, the estimated atmosphere by using 12 fields has even a value 1.0 which is very good. This is a motivating result and if we regard the fact, that in these cases no changes in the frequency space were done, the results could even more improved. Beside that, it has to be stressed at this point that the tested samples are not fully independent from each other. If we test the estimated atmosphere using monthly *masks* in the second point, we receive a p-value of 0.1587 which is very low, but a look at the h-value shows that it is 0, so the test is accepted. Thus, it can be said that the results here can be seen as a good and motivating result, but further tests should be done to prove these results. Now we have done nearly all comparisons, but during the work, we also learned a lot about the frequency space. So, the question occurs, what would be the result, if we calculate the already known correlation coefficients for the estimated atmospheres and the given atmosphere through the time? The point, why we have not done it earlier is, that both samples contain a yearly signal, which was also eliminated in the current significance test. This would also contribute to the result of the correlation calculation and would change the result. However, now we know how to eliminate it and in the following chapter, we will conduct this test.

## 3.7 A last correlation test

As already said in the previous chapter, this chapter has the aim to analyse the correlation through the time, to test, if there are also similarities. But we first have to go again into the frequency domain to eliminate the yearly frequency. This could be easily imagined because both signals contain a yearly signal (because they have a one year duration). Therefore, if we test two samples through the time and both have one signal identical, it is clear that there is a high correlation between them. In the previous tests, we had a result for every day, so it was a comparison between two samples at a fix time, to analyse the correlation between the two places. Now the opposite happens and the place is fix while the time is changing. Because of that, the year frequency has to be eliminated in the frequency domain. After this is done, the first step is, to compare the four points of the daily solutions, which were already used in the previous chapters and calculate the correlation coefficients with the given atmosphere. How the correlation is calculated can be seen in equation 2.24, which is used for the current calculation by using the 'corrcoef' function, where the result is the following matrix:

$$R = \begin{pmatrix} \rho(A,A) & \rho(A,B) \\ \rho(B,A) & \rho(B,B) \end{pmatrix} \quad (3.16)$$

Thus, regarding the matrix in equation 3.16, we need the element at the upper right position to receive the correlation between the two vectors, which gives us the following result:

***Table 3.19:*** *Results of the correlation calculation of the daily solution and the given atmosphere, after the eliminating of the year frequency*

| Point coordinates | $\rho(A,B)$ |
|---|---|
| $(63°, 36°)$ | 0.3906 |
| $(225°, 25.71°)$ | 0.1970 |
| $(36°, 85.71°)$ | 0.1850 |
| $(297°, 120°)$ | 0.1432 |

Table 3.19 shows a bad result, because only the upper point $(63°, 36°)$ have a little higher value than 0.3. All other numbers are between 0.1 and 0.2, which is very bad.
This could also be seen in the calculation of the similar median in the comparison of the two signals in the frequency domain in figure 3.48. In that figure, the two signals of the atmosphere and the daily solutions have several similar pattern, which support the assumption that the actual result is correct and there is a high correlation, which is at least confirmed with the significance test in table 3.17. In contrast, the lower two points show a smaller correlation are also visible in the frequency domain in figure 3.49, here it also can be seen, why there is a very low correlation. The reason is that the daily solutions have a stronger signal, which is probably not from the atmosphere but caused by different effects. The test of equal standard deviation also approves this assumption in table 3.16, in which the two samples even do not have the same standard deviation. Here, it could be again helpful to conduct further analysis which would exceed this work. Now, also the estimated atmospheres can be tested:

***Table 3.20:*** *Results of the calculation of the correlation coefficient of the estimated and the given atmosphere*

| Point coordinates | Atm. using 11 day *masks* $\rho(A,B)$ | Atm. using monthly *masks* $\rho(A,B)$ | Atm. using 1 field $\rho(A,B)$ | Atm. using 2 fields $\rho(A,B)$ | Atm. using 4 fields $\rho(A,B)$ | Atm. using 12 fields $\rho(A,B)$ |
|---|---|---|---|---|---|---|
| $(63°, 36°)$ | 0.8854 | 0.7107 | 0.9305 | 0.8820 | 0.9069 | 0.8863 |
| $(225°, 25.71°)$ | 0.1984 | 0.6369 | 0.8299 | 0.7887 | 0.8041 | 0.7572 |
| $(36°, 85.71°)$ | 0.0796 | 0.0148 | 0.4012 | 0.3454 | 0.2772 | 0.3761 |
| $(297°, 120°)$ | -0.2979 | 0.3641 | 0.4826 | 0.3991 | 0.5353 | 0.4038 |

In this table 3.20 it can be seen that there are mixed results. A look at the estimated atmospheres where we used *masks* shows that only the first point $(63°, 36°)$ has a good result.

The second point is in the estimated atmosphere with 11 day *masks* only at 0.1984 but it still has a good value in the other method, where the monthly *masks* are used, there it has 0.6369 which is still good. That is an interesting fact, because in the previous test, also the estimated atmosphere with 11 day *masks* showed a bad result at the second point, while the estimated atmosphere using monthly *masks* had a good result here. The last two points have a very low correlation, as it can be seen in the previous tests, too. In the previous analyse in the frequency domain in figure 3.49 the reason can be seen in the strong daily solutions, but the reason for the very low correlation should be analysed closer in further tests. The estimated atmospheres, where the given atmosphere is used as start fields and the daily solutions are added to create the estimated atmosphere has better results, which is interesting, because previous tests have shown that they have lesser similarities with the given atmosphere, for example in the test on a similar standard deviation in table 3.16. However, in this case the correlation coefficients of the first two days show very high results, in some cases even higher values than 0.9. But also the lower two points are still in a low level. The test also shows that the usage of one start field has one of the best results, which is surprising, because in the previous tests it had not offered the best results. The second best result has the usage of four start fields and the usage of two and twelve start fields have nearly the same results. Thus, it can be assumed that in this case the method by using the start fields has the best results, but with regard to the previous test, also the estimated atmospheres by using *masks* have good results for the two upper points.

This correlation test is a good way to complete the series of tests, because it confirms the previous results and strengthens the assumption that both methods provide a good result to estimate the atmosphere. The question, why the lower two points (36°, 85.71°) and (297°, 120°) show a lower correlation in all tests has been discussed during the analysis of the results. However, with this test it could be shown that there is also a correlation between the estimated atmospheres and the given atmosphere throughout the time.

## 3.8 Conclusion of the results

The primary goal of this tests was to show that there are significant signals in the daily solutions from the atmosphere. This could be shown clearly by the reconstruction of the atmosphere for a whole year only by using the daily solutions. However, it was also a goal, to test the estimated atmospheres to see, how many similarities they have with the given atmosphere. At this point, the usage of *masks*, as done in the first test, have offered the best results, but this requires atmospheric models for each month. If they are not available, another good method is, to use one or more atmospheric fields as start fields and add the differences of two daily solutions to them. On this way, high similarities also can be reached, but in all cases the peaks of the factor calculation appear at the same days, so this should be analysed closer. Later, we also had a short look in the frequency domain and there it could be seen, that mostly the higher frequencies make a problem, because there, the daily solutions show more energy compared to the given atmosphere. Here, a low pass filter could improve the result. At last, we were able to consolidate the results with the significance test and a second correlation test through the time. There, it was more prominent that the two lower points which are tested show a worse result compared to the upper two points. The reason could be seen in the frequency domain, because the given atmosphere has a lower energy compared to the daily solutions. But the reason, why there is such a big difference should be analysed in further tests. However, all tests have shown that the estimated atmospheres have a high correlation with the atmosphere and even the daily solution has a significant correlation to the atmosphere.

# Chapter 4

# Conclusion of the results

## 4.1 Overview of the results

Now, we come to the end of our excursion to the daily solutions and gravity fields. Our aim was to estimate atmospheric signals by using the daily solutions. At the beginning, we learned a lot about the Daily Wiese Solutions and how to calculate them, then we started to compare the daily solutions with the atmosphere and we could already see some similarities, which gave us a first sign that we are on the right way. The next chapter started with the estimation of the atmosphere, where we made some surprising discoveries by testing two methods, to estimate the atmosphere. The first possibility is to calculate *masks*, which can be used to create the atmosphere, the other method is to add the differences of two daily solutions to a given atmosphere, to reconstruct the atmosphere by only using the daily solutions with a few atmospheric fields. At the end, it can be said that we found two good methods, both with different advantages and disadvantages. The first method with the *masks* offers the best results to estimate the atmosphere, it is the method of choice which can be recommended if someone wants to use the daily solutions to estimate the atmosphere. The problem is that we need a model of the atmosphere for the whole year. The other method has a lower correlation with the given atmosphere, but the advantage is that no model of the whole atmosphere is needed. Here, we have again the peaks, which should be analysed closer. But beside the lower results, if we regard the fact that we talk about the atmosphere, which is not reconstructed completely, but to a high correlation by using the daily solutions, we realise that the primary goal to estimate atmospheric signals in the daily solutions, is achieved. After these tests, we took a short look into the frequency domain to see, how the daily solutions and the given atmosphere match, and we could see that there are already similarities. We also could see that the high frequencies show more energy compared to the given atmosphere. Here, it is interesting because the tested lower two points show fewer similarities as the upper two points, because there the energy is much stronger compared to the atmosphere. So there is also one point, which can be improved in further tests. At the end, a significance test and also a second correlation test were done to see, if the estimated atmospheres match significant with the given atmosphere. These results show that there are already significant matches between them and when we regard the mentioned points above, we see that this results can already be improved. It clearly can be said that there are atmospheric signals in the daily solutions. The added aim, to reconstruct the atmosphere is also achieved, but as mentioned there could be some further improvements which would go beyond this thesis.

These results can be used as a basis to start with further analyses which would extend the limit of this thesis. One possibility would be more analyses in the frequency domain, to see, how the variations we saw in the spacial domain appear there and if there a filter perhaps can be used to increase the correlation. Here, it would be also very interesting to analyse the mentioned peaks, which we could not analyse in the needed detail to find out, why they appear. In the frequency space we can also make further comparisons between the daily solutions and the atmospheres but also between the estimated atmospheres and the given atmospheres. The other big point which could be analysed closer are the big variations between the upper two points and the lower two points. There, it would be also helpful to do closer analyses.

This thesis can be seen as the beginning of the analysis of atmospheric signals contained in the measured gravity field. Hopefully, further research will enable the weather forecast to use gravity science for their work.

# Bibliography

(n.d.).

Angel, J. (n.d.), 'Record rainstorm and flood of july 17-18, 1996'.
**URL:** *http://www.weather.gov/lot/1996Jul1718_rain*

*Atmospheric pressure, winds, and circulation patterns, p. 113 - 137* (n.d.).
**URL:** *http://www.cengage.com/resource_uploads/downloads/0495555061_137182.pdf*

Bender, P.L., W. D. N. R. (2008), Proceedings of the third international symposium on formation flying, missions and technologies, *in* 'A possible dual-GRACE mission with 90 degree and 63 degree inclination orbits', ESA/ESTEC, Noordwijk, The Netherlands, pp. pp. 1 – 6.

*Cold May Of 1996 /Warm May Of 1992* (n.d.).
**URL:** *http://www.ukweatherworld.co.uk/forum/index.php?/topic/44039-cold-may-of-1996-warm-may-of-1992/*

Dobslaw, H., B. I. D. R. F. E. K. V. K. J. S. I. (2014), Updating esa' earth system model for gravity mission simulation studies: 1. model description and validation, Scientific technical report 14/07, GFZ German Research Centre for Geosciences.

*Historical Weather For 1996 in Lucknow, India* (n.d.).
**URL:** *https://weatherspark.com/history/33937/1996/Lucknow-Uttar-Pradesh-India*

*Historical Weather For 1996 in New Delhi, India* (n.d.).
**URL:** *https://weatherspark.com/history/33934/1996/New-Delhi-India*

Loomis, B. (2009), Simulation study of a follow-on gravity mission to GRACE, PhD Dissertation, PhD thesis, Colorado Center for Astrodynamics Research, University of Boulder, Boulder, CO, USA.

Mat (n.d.*a*), *fft - Fast Fourier transform*.
**URL:** *https://de.mathworks.com/help/matlab/ref/fft.html*

Mat (n.d.*b*), *Fourier Transforms*.
**URL:** *https://de.mathworks.com/help/matlab/math/fourier-transforms.html*

Mat (n.d.*c*), *ranksum*.
**URL:** *https://de.mathworks.com/help/stats/ranksum.html*

Mat (n.d.*d*), *vartest2 - Two-sample F-test for equal variances*.
**URL:** *https://de.mathworks.com/help/stats/vartest2.html*

*National Weather Service, Storm Data and Unusual Weather Phenomena* (n.d.).
**URL:** `http://www.weather.gov/media/bis/sd1996May.pdf`

Palvis, D.E., D. C. M. J. (2010), *GEODYN operations manual. Contactor Report*, SGT, Inc., Greenbelt, MD.

Rowlands, D.D., R. R. C. L. F. (2002), 'Short-arc analysis of intersatellite tracking datain a gravity mapping mission', *J. Geod.* **76**, 307–316.

*SHBundle: http://www.gis.uni-stuttgart.de/research/projects/bundles/* (n.d.).

Sneeuw, N. (2006), *Physical Geodesy, Lecture Note of the cours*.

Sneeuw, N., W. M. A. M. R. M. D. (2015), Shbundle, Technical report, Institute of Geodesy.

Tapley, B.D., S. B. B. G. (2004b), 'Statistical orbit determination', *Elvsevier Inc., San Diego* .

Ullman, R. (1997), 'Solve program mathematical formulation', *Rep. HSTX-G and G-9201* .

Wahr, J., M. M. B. F. (1998), *JOURNAL OF GEOPHYSICAL RESEARCH* **B12**, 30, 205–30,22.

Wiese, D., V. P. N. R. (2011), 'Estimating low resolution gravity fields at short time intervals to reduce temporal aliasing errors', *ScienceDirect, Advances in Space Research 48 (2011)* **48**(2011 COSPAR. Published by Elsevier Ltd. All rights reserved. doi:10.1016/j.asr.2011.05.027), 1094–1107.

*Wikipedia: Alias-Effekt* (n.d.).
**URL:** `https://de.wikipedia.org/wiki/Alias-Effekt`

*Wikipedia: Gravity Recovery and Climate Experiment* (n.d.).
**URL:** `https://en.wikipedia.org/wiki/Gravity_Recovery_and_Climate_Experiment`

*Wikipedia: Korrelationskoeffizient* (n.d.).
**URL:** `https://de.wikipedia.org/wiki/Korrelationskoeffizient`

*Wikipedia: Luftdruck* (n.d.).
**URL:** `https://de.wikipedia.org/wiki/Luftdruck`

*Wikipedia: Südhalbkugel* (n.d.).
**URL:** `https://de.wikipedia.org/wiki/S\%C3\%BCdhalbkugel`

*Wikipedia: Zero drag satellite* (n.d.).
**URL:** `https://en.wikipedia.org/wiki/Zero-drag_satellite`

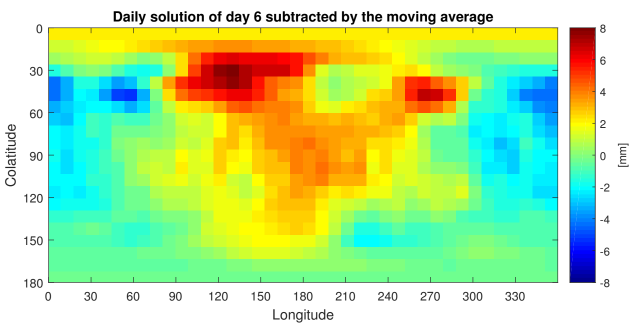
*Wilcoxon-Mann-Whitney-Test* (n.d.).
**URL:** `https://de.wikipedia.org/wiki/Wilcoxon-Mann-Whitney-Test`

Zenner, L. (2005), *zfv - Zeitschrift für Geodäsie, Geoinformation und Landmanagemen* **Vol.132, Nr. 3**(DOI: 10.1111/j.1365-246X.2010.04669.x), pp 147–157(2007).

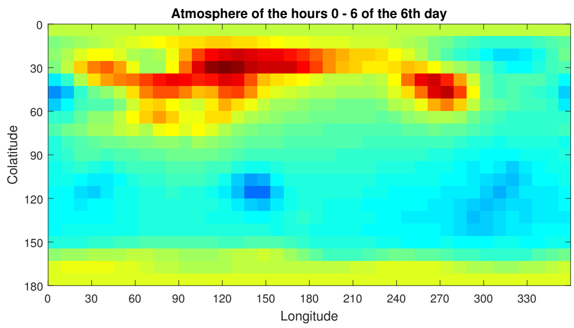
# Appendix A

## A.1 Correlation between the epochs of the atmosphere and the daily solution

At the beginning, the first epoch from 0h to 6h will be shown. In the following comparison, the daily solution of the 6th day is compared to the atmosphere of the hour 0 - 6 from the same day:




(a) Daily solution reduced by the Moving Average




(b) Atmosphere from the hours 0 - 6

***Figure A.1:*** *Comparison of the daily solution and the first epoch of the atmosphere*

Here, the strong positive pattern in the middle upper area can be seen in both fields. Also, a stronger positive pattern in the upper right area appear in both fields and also on the upper left side a small negative area can be now clear seen in both fields. The stronger positive area of the atmosphere in the lower frame can not be seen in the daily solution. In the opposite, the daily solution shows a positive pattern in the middle of the field, where the atmosphere has a weak negative pattern. So the two striking positive areas in the atmosphere can be seen in the daily solution but beside that, there are not many similar pattern. In the next step, the correlation coefficients will be shown. At first, the normal correlation coefficient will be calculated:

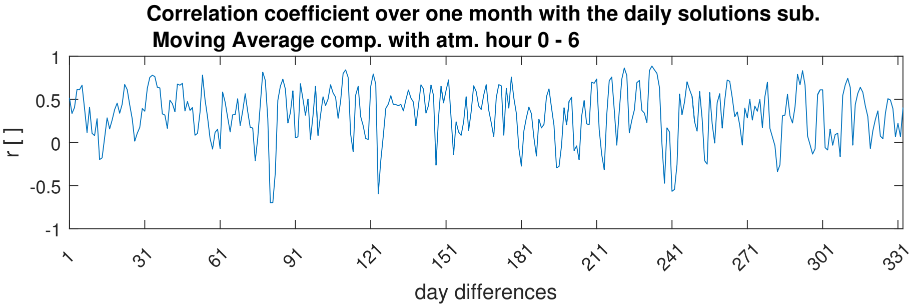



*Figure A.2: The correlation coefficient over one year for the daily solution and the first atmospheric epoch from 0 - 6h*

Here, 115 coefficients are over 0.5, which is not very high, compared to the previous tests. The mean value is 0.3367 which is not very high. Now, the correlation will be calculated, when the spherical form of the earth is taken into account by multiply the fields with $sin\theta$:

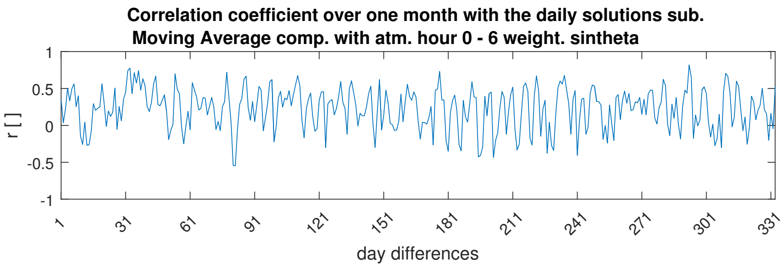



*Figure A.3: The correlation coefficient over one year for the daily solution and the first atmospheric epoch from 0 - 6h, with the multiplication of sin theta*

In this case, 57 coefficients are higher than 0.5 and a mean coefficient of 0.2291, which is both lower than in the previous calculation. In the next step, the correlation with the daily solutions, where the zonal coefficients are subtracted is calculated:

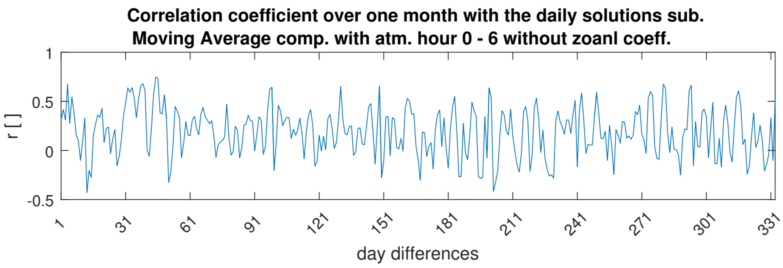



*Figure A.4: The correlation coefficient over one year for the daily solution and the first atmospheric epoch from 0 - 6h without the zonal coefficients*

Here, only 41 coefficients are higher than 0.5, which is again a lower value, the average coefficient is 0.1951, which is also very low. At least, the coefficients without the zonal coefficients also with regarding the spherical form of the earth, with the $sin\theta$, can be calculated:

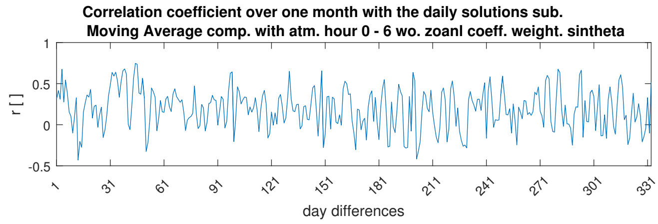
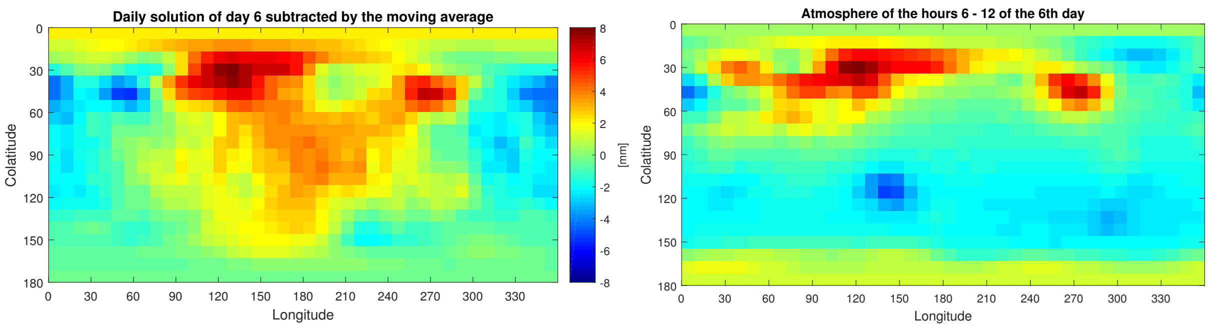



***Figure A.5:*** *The correlation coefficient over one year for the daily solution and the first atmospheric epoch from 0 - 6h without the zonal coefficients, by weighting the fields with sin theta*

After this calculation, again 41 coefficients are higher than 0.5, with a mean coefficient of 0.1951, which are similar values compared to the previous test. Here we can say that there is no real improvement. Now, we have a look at the atmosphere between the hours 6 - 12, at first again a comparison between the atmospheric field and the daily solution of the same day:




(a) Daily solution reduced by the Moving Average (b) Atmosphere from the hours 6 - 12

***Figure A.6:*** *Comparison of the daily solution and the second epoch of the atmosphere*

Here, nearly the same similarities like in the previous comparison. The most prominent change is that now the power of the positive pattern in the upper left area of the atmospheric field decrease. In the next step, we can have a look how the changes affect the calculation of the correlation coefficients:

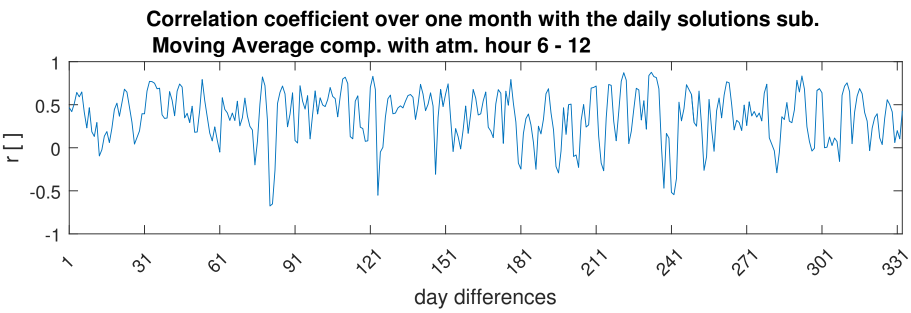



***Figure A.7:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 6 - 12h*

This case shows that there are 131 coefficients over 0.5, compared to the previous calculated 115 coefficients it is an improvement and the mean value is 0.3654 which is also higher compared to the previous epoch. As done before we can have a look at the coefficients after the multiplication of the $sin\theta$ values:

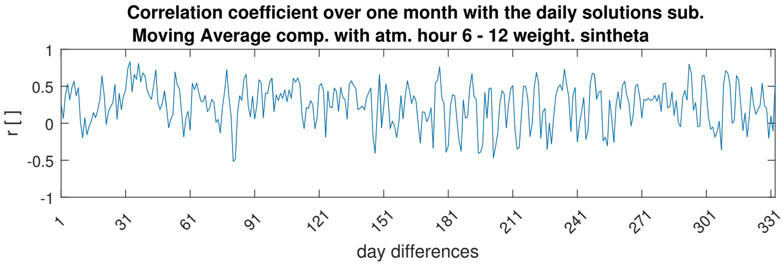



***Figure A.8:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 6 - 12h, with the multiplication of sin theta*

Also here is now a lower number of coefficients over 0.5, only 74 have such a value, compared to the last coefficients which are calculated with $sin\theta$ there is an improvement. The mean correlation coefficient is 0.2542, which is also an improvement, to the similar calculation. In the next step, the zonal coefficients of the daily solutions will be removed and the correlation is calculated again:

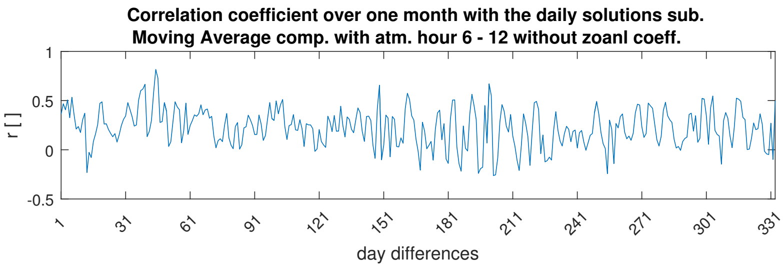



***Figure A.9:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 6 - 12h without the zonal coefficients*

This result shows a worse result, because only 26 coefficients are higher than 0.5, which is lower than the previous one. In the opposite, the mean coefficient is 0.2284, which is an improvement, compared to the previous one. Now, the same fields will be calculated, with the multiplication of the $sin\theta$ values:

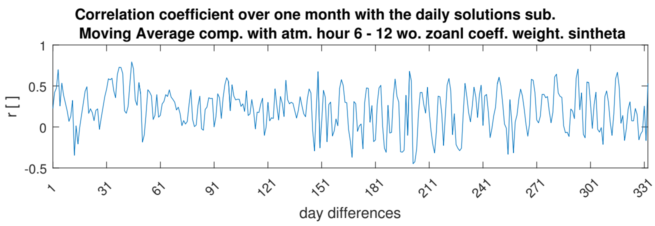
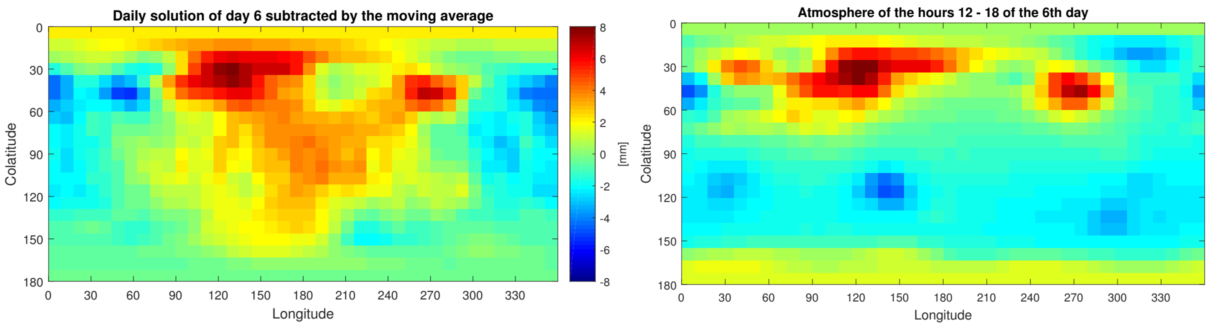



***Figure A.10:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 6 - 12h without the zonal coefficients, by weighting the fields with sin theta*

Here, 49 coefficients higher than the value of 0.5, which is an improvement, but still low. The mean coefficient is 0.2177, so only a bit lower as the one without the multiplication of the $sin\theta$ values. It can be said that the using of the second epoch is an improvement. In the following, we test, if the third epoch provides again an improvement:




(a) Daily solution reduced by the Moving Average (b) Atmosphere from the hours 12 - 18

***Figure A.11:*** *Comparison of the daily solution and the third epoch of the atmosphere*

The comparison of the two fields shows that the positive area in the atmospheric field decrease so that it match better with the stronger positive pattern of the daily solution. Therefore we can test, if this change improves the result of the correlation coefficients:

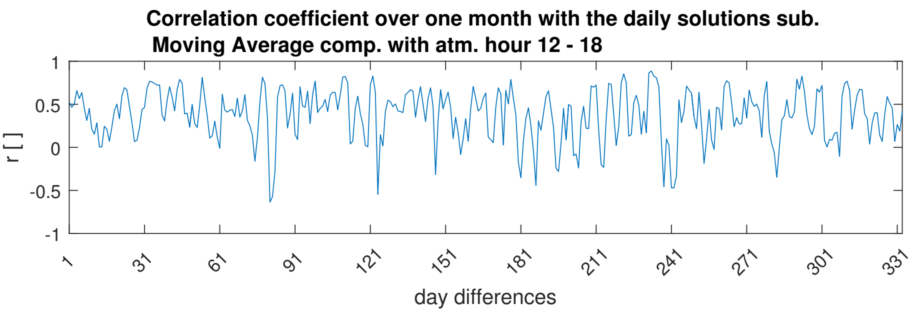



***Figure A.12:*** *The correlation coefficient over one year for the daily solution and the third atmospheric epoch from 12 - 18h*

Here, we have no improvement, because with 130 coefficients over 0.5 it is a bit lower compared to the previous test. The mean coefficient over one year is 0.3779 which is an improvement. In the following, the correlation with the using of the $sin\theta$ values is shown:

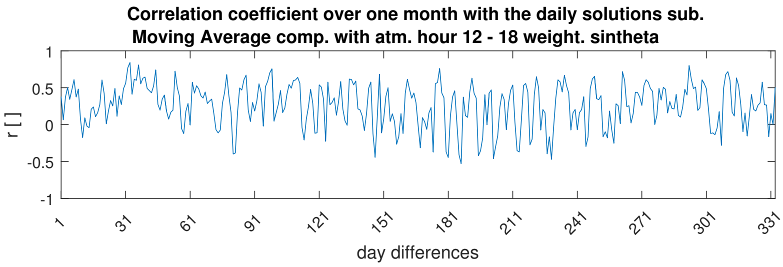



***Figure A.13:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 12 - 18h, with the multiplication of sin theta*

In this case, the number of coefficients increase, with 80 coefficients higher than 0.5, and a mean correlation coefficient of 0.2646. During the next step we can have a look at the correlation coefficients, after the removing of the zonal coefficients:

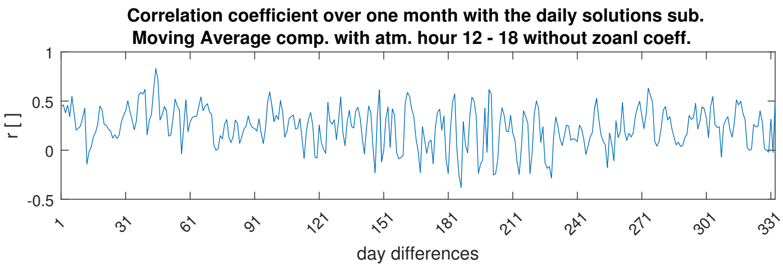



***Figure A.14:*** *The correlation coefficient over one year for the daily solution and the third atmospheric epoch from 12 - 18h without the zonal coefficients*

This analysis shows that the number is a bit higher as the previous ones, with only 31 coefficients over the value of 0.5 and a mean coefficient over one year with 0.2378, which is also an increase in the value. In the next step, the coefficients, calculated with the multiplication of the $sin\theta$ values, will be shown:

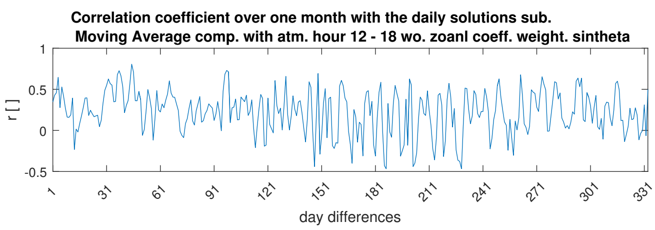
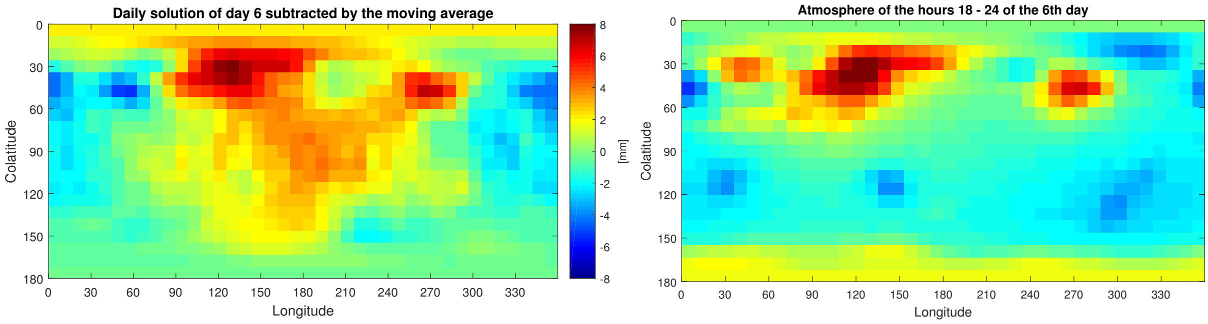



***Figure A.15:*** *The correlation coefficient over one year for the daily solution and the third atmospheric epoch from 12 - 18h without the zonal coefficients, by weighting the fields with sin theta*

With 51 values higher than 0.5, there is again an increase in the coefficients and the mean value is 0.2265, which is again higher than the previous ones. At least, we test the last epoch from the hours 18 - 24:




(a) Daily solution reduced by the Moving Average (b) Atmosphere from the hours 18 - 24

***Figure A.16:*** *Comparison of the daily solution and the fourth epoch of the atmosphere*

The comparison shows that the positive pattern in the upper left of the atmosphere decrease again and also the negative pattern at the upper right side appear stronger, which match with a negative pattern in the daily solution. So now it can be tested, if there is an improvement:

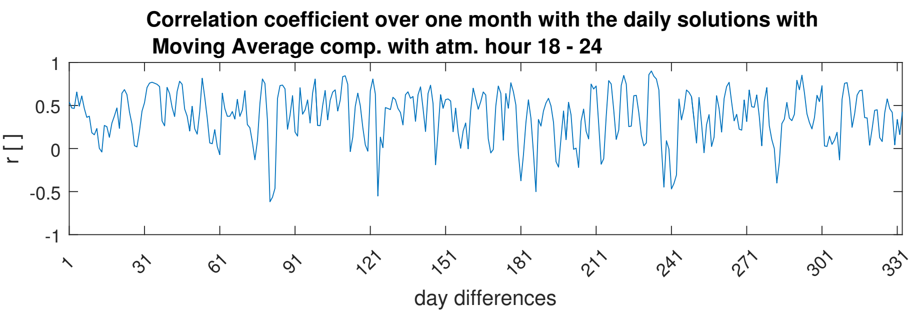



***Figure A.17:*** *The correlation coefficient over one year for the daily solution and the 4th atmospheric epoch from 18 - 24h*

But with 130 coefficients over 0.5, we have the same result compared to the previous test. The mean coefficient over one year is 0.3766, which is a bit lower as the previous one. In the following test, the result can be seen, if the correlation coefficients with $sin\theta$ is calculated:

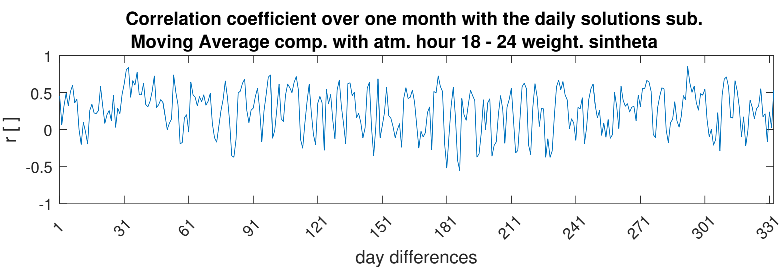



***Figure A.18:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 18 - 24h, with the multiplication of sin theta*

Here, now 84 coefficients are higher than 0.5, which is higher as the previous 80. The mean value is 0.2573, which is lower as the previous calculation. Now, we test the result with the removing of the zonal coefficients:

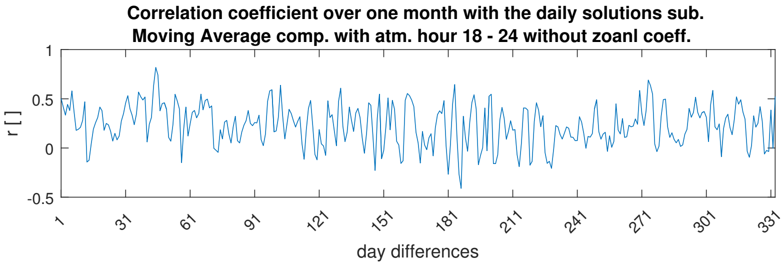



*Figure A.19:* *The correlation coefficient over one year for the daily solution and the 4th atmospheric epoch from 18 - 24h without the zonal coefficients*

In this case, 34 coefficients are higher than 0.5, which is again an improvement to the previous coefficients. A look at the mean coefficient over one year shows that it is 0.2356, which is lower than the previous one. At least, we can have a look at the correlation, which respects the spherical form of the earth:

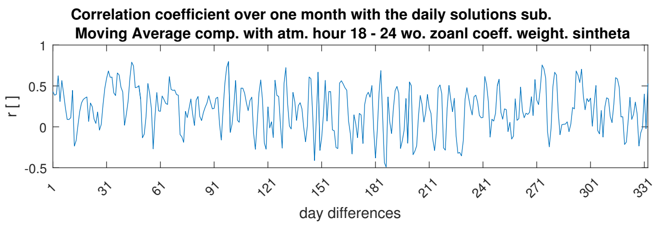
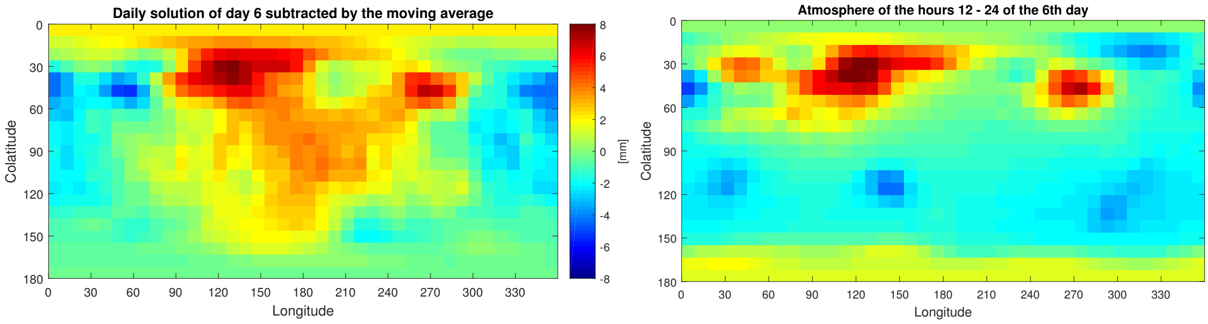



*Figure A.20:* *The correlation coefficient over one year for the daily solution and the 4th atmospheric epoch from 18 - 24h without the zonal coefficients, by weighting the fields with sin theta*

Also here a trend can be seen, because there is again an improvement with 58 coefficients over 0.5 we can see an improvement but the mean coefficient is lower than the previous one with 0.2193. As a conclusion it can be said that the last two epochs show the best results, the third epoch has the highest mean values in the most comparisons and the last epoch has the most coefficients higher than 0.5. So it could be very interesting, if there is an improvement, if we calculate a mean field of these two fields of the atmosphere and compare it with the daily solutions:




(a) Daily solution reduced by the Moving Average

(b) Atmosphere from the hours 12 - 24

***Figure A.21:*** *Comparison of the daily solution and the third and fourth epochs of the atmosphere*

Because it is the sum of the last two compared fields, here the changes can be seen together. The positive pattern of the atmosphere on the left side is smaller and match with the positive pattern in the same area of the daily solution. Also the small positive pattern on the right side can be seen in both fields. But as seen before, all similarities are only in the upper area of the fields. Now we can have a look, if there is an improvement in the calculated correlation:

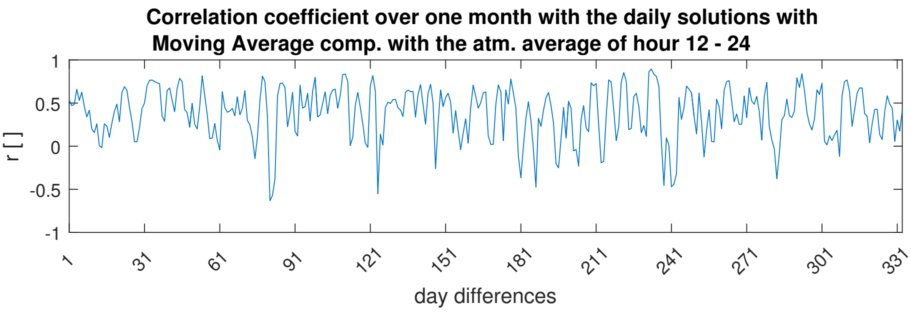



***Figure A.22:*** *The correlation coefficient over one year for the daily solution and the sum of the 3rd and 4th atmospheric epoch*

The result shows that with 131 coefficients higher than 0.5, the number is similar to the other highest number until now, which can be seen in epoch 2. The mean coefficient over one year is with 0.3789 the highest one. Now, we see the result, after the multiplication with the $sin\theta$ values:

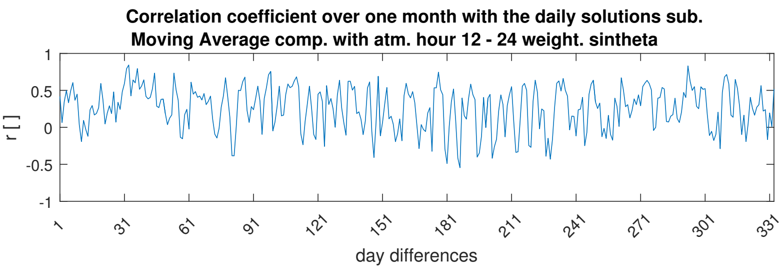



***Figure A.23:*** *The correlation coefficient over one year for the daily solution and the second atmospheric epoch from 12 - 24h, with the multiplication of sin theta*

These calculation shows that there are 88 coefficients higher than 0.5, which is the highest number, but the mean correlation coefficient is with 0.2633 a bit lower than the mean value of epoch 3. In the next step, we test the result after the re-movement of the zonal coefficients:

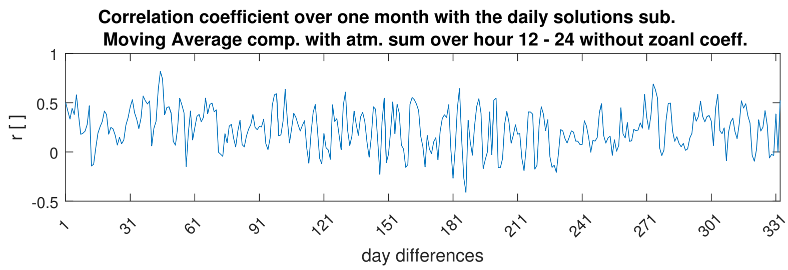



*Figure A.24:* *The correlation coefficient over one year for the daily solution and of the 3rd and 4th atmospheric epoch without the zonal coefficients*

With 34 coefficients higher than 0.5, we have no real improvement to the previous fields because it is similar to the number of the fourth epoch. The mean value is 0.2378, which the highest value in the comparison. At least, we can test the correlation, after the fields are multiplied with *sinθ*:

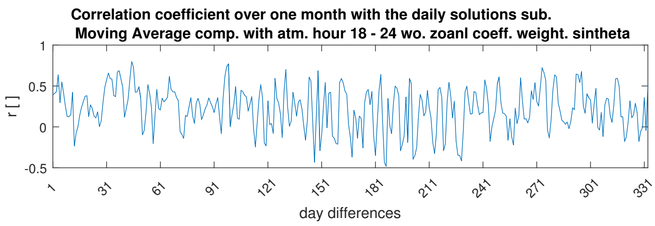



*Figure A.25:* *The correlation coefficient over one year for the daily solution and of the 3rd and 4th atmospheric epoch without the zonal coefficients, by weighting the fields with sin theta*

In this last test, there are 53 coefficients higher than 0.5, which is the second highest number of coefficients, the mean value in this calculation is 0.2252 and also the second highest value.