Radiosonde Tracking and Calibration Studies with the In-ice Phased Array of the Radio Neutrino Observatory Greenland (RNO-G)

Bachelorarbeit aus der Physik

Vorgelegt von

Carolin Klein

21.08.2023

Friedrich-Alexander-Universität Erlangen-Nürnberg

Betreuer:
PD Dr. Robert Lahmann

Abstract

The Radio Neutrino Observatory Greenland is currently under construction. During the construction phase of an experiment, it is important to check for inconsistencies and potential issues in the functionality of the detector. This could lead to misinterpretation of the results when data taken by the finalized detector setup is analyzed. For this purpose, a reconstruction of the track of a radiosonde using the measured signals in the in-ice phased arrays of the deployed stations is performed. The angle of the radiosonde signal is determined using the delay-and-sum beamforming technique. Based on the observations, a method is proposed to detect and characterize calibration issues related to cable delays and antenna spacing of the phased array. Additionally, an alternative possibility to detect inaccuracies in the antenna spacing is proposed taking the refractive index of ice into account. The reconstruction of the radiosonde position is performed using a χ^2 minimization. For some of the investigated stations, inconsistencies in the calibration are observed. However, the radiosonde tracking in y- and z-direction, respectively northing and altitude, provides good results. For the observed deviations in the reconstruction of the x-coordinate (easting), the arrangement of the stations considered and an angle deviation in one station are identified as the source.

Contents

1 Introduction

The Universe has occupied mankind since time immemorial. Stars were observed, their movements documented and analyzed. People wanted to find out how our world is structured. But it is not always helpful to observe the big things. Sometimes the small things give us information about the whole. For example, more than a billion particles of cosmic origin arrive on Earth every second. Observing them can tell us how the universe is formed.

Especially neutrinos are an important part of the so-called multi-messenger astronomy. Neutrinos are elementary particles. Due to their non-existing electric charge, they are not deflected on their way from the source to Earth. Therefore, they point to the source in which they were produced. However, the small interaction cross sections of neutrinos makes their detection difficult [1].

The Radio Neutrino Observatory Greenland, or RNO-G, is being built to detect high-energy neutrinos via radio waves using the Askaryan effect and to study their origin in more detail. Since 2022 7 stations have been operational. 3 were deployed in 2021, 4 in 2022. For the Askaryan effect to take place, a medium with a higher refractive index than air is necessary. Therefore, the antennas of the RNO-G stations are installed in ice. The heart of the RNO-G stations is formed by an in-ice phased array that allows to reconstruct the incident angle of a radio wave [1].

During the construction phase of an experiment, it is important to check for inconsistencies and potential issues in the functionality of the detector. This could lead to misinterpretation of the results when data taken by the finalized detector setup is analyzed.

In some cases, actual sources of interference can be used, e.g. for calibration or calibration verification. For RNO-G, a weather balloon equipped with a radiosonde serves as such a source. The weather balloon is launched twice a day as part of the Integrated Characterization of Energy, Clouds, Atmospheric state, and Precipitation at Summit (ICECAPS) project, which studies the interaction between atmospheric processes and the Greenland ice sheet [2]. The radiosonde of the balloon emits radio waves for data transmission allowing to track the GPS position of the radiosonde, respectively the balloon.

The goal of this thesis is to verify if a reconstruction of the track of a radiosonde using the measured signals in the in-ice phased arrays of the deployed stations of RNO-G is possible. The first step is to determine the angle at the phased array of the radio wave transmitted by the balloon. This is done using the delay-and-sum beamforming technique [3]. During this procedure attention is paid to inaccuracies in calibration, e.g. to possible inaccurate cable delays. Based on the observations, a method is proposed to detect and characterize calibration issues related to cable delays and antenna spacing of the phased array. Additionally, an alternative possibility to detect inaccuracies in the antenna spacing is proposed taking the refractive index of ice into account. Finally, the reconstruction of the balloon position is performed using a χ^2 minimization.

2 Theoretical Background

This section provides a brief overview of the theoretical background of this thesis. First, it is explained what neutrinos are. Then, the radio detection of neutrinos is described, including the setup of the Radio Neutrino Observatory Greenland. Afterwards, weather balloons and radiosondes are introduced and coordinate systems are discussed.

2.1 Neutrinos

Neutrinos are fundamental particles in the Standard Model of particle physics. This model describes the currently known fundamental particles in our universe and the interactions between them.

In 1930, Pauli postulated a new particle to explain the three-body nature found in the electron spectrum of beta decay ($\bar{\nu}_\mathrm{e} + \mathrm{p} \rightarrow \mathrm{e}^+ + \mathrm{n}$). To avoid perturbing the spin statistics of the decay, the new particle had to be a particle with a spin of 1/2. It also had to be electrically neutral and carry very little mass. These properties led Pauli to the following statement [4]:

> "I have done a terrible thing, I have postulated a particle that cannot be detected."

This postulated particle is now called a neutrino which is, contrary to Pauli's proclamation, not impossible to detect. However, due to the small interaction cross sections of neutrinos and the absence of electric charge, detection is indeed quite complicated.

The first experiment that actually detected neutrinos (Savannah River Experiment) used a sandwich detector consisting of scintillation detectors and tanks filled with a mixture of water and cadmium. The goal of the experiment was to detect anti-electron neutrinos, which were produced by beta decay in a nuclear power plant. When one of these anti-electron neutrinos underwent inverse beta decay ($\bar{\nu}_\mathrm{e} + \mathrm{p} \rightarrow \mathrm{e}^+ + \mathrm{n}$) in one of the water tanks, a positron was created that underwent annihilation, producing a photon pair that would then induce a signal in the scintillator. The neutron would be slowed down and eventually captured by the cadmium, again producing a signal in the detector. A separation of the pulses was a sign for the inverse beta decay and therefore for the existence of the neutrino [5].

The Savannah River Experiment detected low energy neutrinos. Much has happened in the field since then. For example, by detecting high-energy neutrinos, IceCube has established a new field of astronomy called neutrino astronomy [6].

2.2 Radio Neutrino Observatory Greenland

In this section, the radio detection of neutrinos is explained. Furthermore, the propagation of radio waves in air and ice is described, as well as the setup of the Radio Neutrino Observatory Greenland (RNO-G).

2.2.1 Radio Detection of Neutrinos

One option to measure high-energy neutrinos is by radio detection. In general, there are two mechanisms by which a neutrino, or more precisely the particle shower resulting from the interaction of a neutrino with a medium, can lead to the generation of radio waves.

The most important effect in air is the so-called geomagnetic emission of the shower. This emission is caused by the transverse separation of the positively and negatively charged particles of the shower by the Lorentz-force (cf. Figure 1). This separation generates a time-varying dipole and thus an electromagnetic wave. In the case of these particle showers, the frequency of the

wave is in the radio range (MHz-GHz) [7].

For media with higher densities, such as ice, a second mechanism is especially important. The shower particles interact with the particles, especially electrons, in the medium. In these interactions, the positrons from the shower annihilate with the electrons in the medium. Other electrons from the medium are extracted from the atoms and become part of the shower, while the positively charged nuclei are left behind. These processes cause a longitudinal charge separation and a time-varying excess of negative charge of about 20% at the shower front (cf. Figure 1). A dipole is created, which in turn emits electromagnetic waves in the radio range. This radiation is called Askaryan radiation [8][1][9]. Similar to Cherenkov radiation, Askaryan radiation is emitted conically along the axis of the shower. At the Cherenkov angle, the signal is strongest, because all electromagnetic waves arrive at the same phase (constructive interference) and therefore coherence is given [8][1].

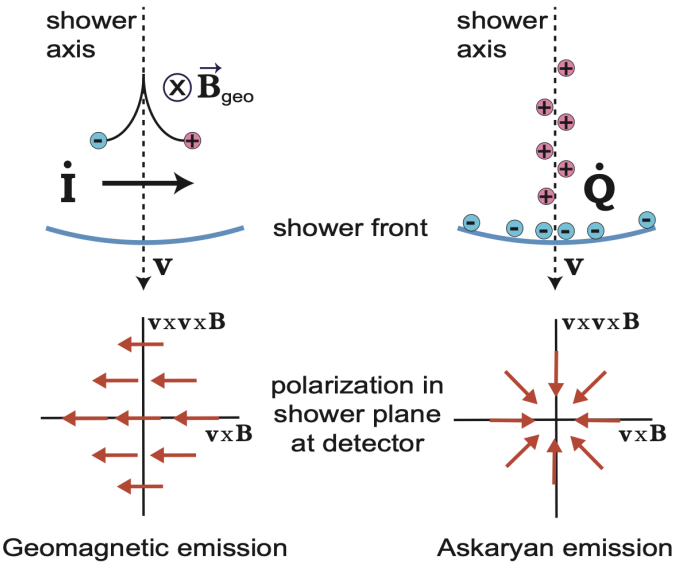


Fig. 1: Main mechanisms for the generation of radio waves from particle showers. Geomagnetic emission important for air (left) and Askaryan emission for denser media, such as ice (right). Figure from Ref. [10].

2.2.2 Radio Wave Propagation in the Atmosphere and Ice

The simplest way to describe electromagnetic radiation is to use geometrical optics, also called ray optics. In a homogeneous medium with a constant refractive index n, the radiation is represented as a straight line, whereas in inhomogeneous media, the ray representing the path of the radiation is curved, depending on the change in refractive index of the medium. The general path $\vec{r}$ of such a ray can be parameterized using the ray equation, also called the eikonal equation [11]:

$$\frac{\mathrm{d}}{\mathrm{ds}}\left(\mathrm{n}(\vec{r})\,\frac{\mathrm{d}\vec{r}}{\mathrm{ds}}\right) = \vec{\nabla}\,\mathrm{n} \tag{1}$$

In this equation s is the so-called arch-length parameter.
Equation 1 follows directly from Fermat's principle, which states that light (or any electromagnetic wave) propagates along the path that can be traveled in the least amount of time [11].

Propagation in the Atmosphere

From Equation 1 it follows that the propagation of radio waves is affected by the refractive index of the medium. In the atmosphere, the refractive index depends on several parameters such as the temperature or the relative humidity of the air, but it is to be noted that changes in the refractive index coming from these variables are small. An approximation for the index of refraction in dependence of the height over the Earth's surface is given by the following formula [12]:

$$\mathrm{n}(h) = 1 + \mathrm{N}_0 \cdot 10^{-6} \cdot \exp\left(\frac{-h}{\mathrm{h}_0}\right)$$

where N_0 is the average value of the extrapolation of the atmospheric refractivity to sea level and h_0 is the scale height given in km. As reference, values of $\mathrm{N}_0 = 315$ and $\mathrm{h}_0 = 7.35\,\mathrm{km}$ can be assumed [12]. More precise values depend again on the climate conditions. Using the reference values the refractive index becomes $\mathrm{n}(h = 0\,\mathrm{km}) = 1.0003$ respectively $\mathrm{n}(h = 30\,\mathrm{km}) = 1.000005$. This demonstrates that the index changes only slowly with the height and is close to one. For this thesis, the refractive index of the atmosphere is therefore assumed to be constant with an index of refraction $\mathrm{n} = 1$.

Propagation in Ice

Ice sheets (for example in Greenland) consist of two regions. The top layer of the sheets consist of loose snow that with increasing depth transforms into glacial ice. This transition zone is called the firn. The second region consists of glacial ice. When looking at the propagation of electromagnetic waves, e.g. radio waves in the case of interest for this thesis, it is important to note that the refractive index of ice can be described using the so called Schytt equation [13]:

$$\mathrm{n}(x, y, z) = 1 + 0.78 \cdot \frac{\rho(x, y, z)}{\rho_0}$$

with ρ being the local density of ice and ρ_0 being the solid ice density with a value of $\rho_0 = 917\,\frac{\mathrm{kg}}{\mathrm{m}^3}$.

For Summit Station (Greenland) several firn density measurements were carried out. Comparison with firn models lead to the following density profile [14]:

$$\rho(z) = \begin{cases} 0.917 - 0.594\exp(-\frac{z}{30.8}), & z \leq 14.9 \\ 0.917 - 0.367\exp(-\frac{z-14.9}{40.5}), & z > 14.9 \end{cases}$$

where z is the vertical coordinate, i.e. the depth in the ice.

A simple ice model for Summit Station can therefore be described by [15]:

$$\mathrm{n}(z) = 1.78 - 0.51\exp\left(\frac{z}{37.25}\right) \tag{2}$$

Transition between Air and Ice

Since ice and air, respectively the atmosphere, have different refractive indices, the transition of radio waves these two media results in refraction. In ray optics, refraction is described by Snell's law [11]:

$$\mathrm{n}_1 \cdot \sin(\theta_1) = \mathrm{n}_2 \cdot \sin(\theta_2) \tag{3}$$

where n_1 and n_2 are the refractive indices of the media between the wave transitions, and θ_1 and θ_2 are the zenith angles of the ray in the different media.

In this thesis, zenith angles in air defined in the interval $[0, \frac{\pi}{2}]$ are considered. Therefore, the maximum angle of radio waves that can be observed in ice with a refractive index of about 1.74 is approximately 35 degrees.

2.2.3 RNO-G Station Setup

The Radio Neutrino Observatory Greenland is an experiment currently under construction (since 2021) at Summit Station in Greenland. Its goal is to detect astrophysical neutrinos with energies from 10 PeV up to EeV using the Askaryan effect. So far, no neutrino in this energy range has been measured using radio detection due to the low neutrino flux [1].

RNO-G is an in-ice experiment. This leads to large effective detector volumes due to the attenuation length of radio waves in ice of about 1 km. In comparison, the attenuation length of optical waves (Cherenkov radiation) in water is only in the order of $50 - 100$ m [16].

When completed, the experiment will consist of 35 detector stations. In 2022, 7 were deployed. Each of them contains two parts, a surface and a deep component. A map of the planned stations and a schematic setup of a single station is depicted in Figure 2.

The surface component consists of nine antennas. In this case log-periodic dipole antennas (LPDAs), which can be used as muon veto. Due to the continuously varying refractive index of the ice, neutrino-generated showers in the ice barely reach the surface and therefore do not reach the LPDAs. Most of the signal they detect is therefore produced by cosmic rays, especially secondary muons [17].

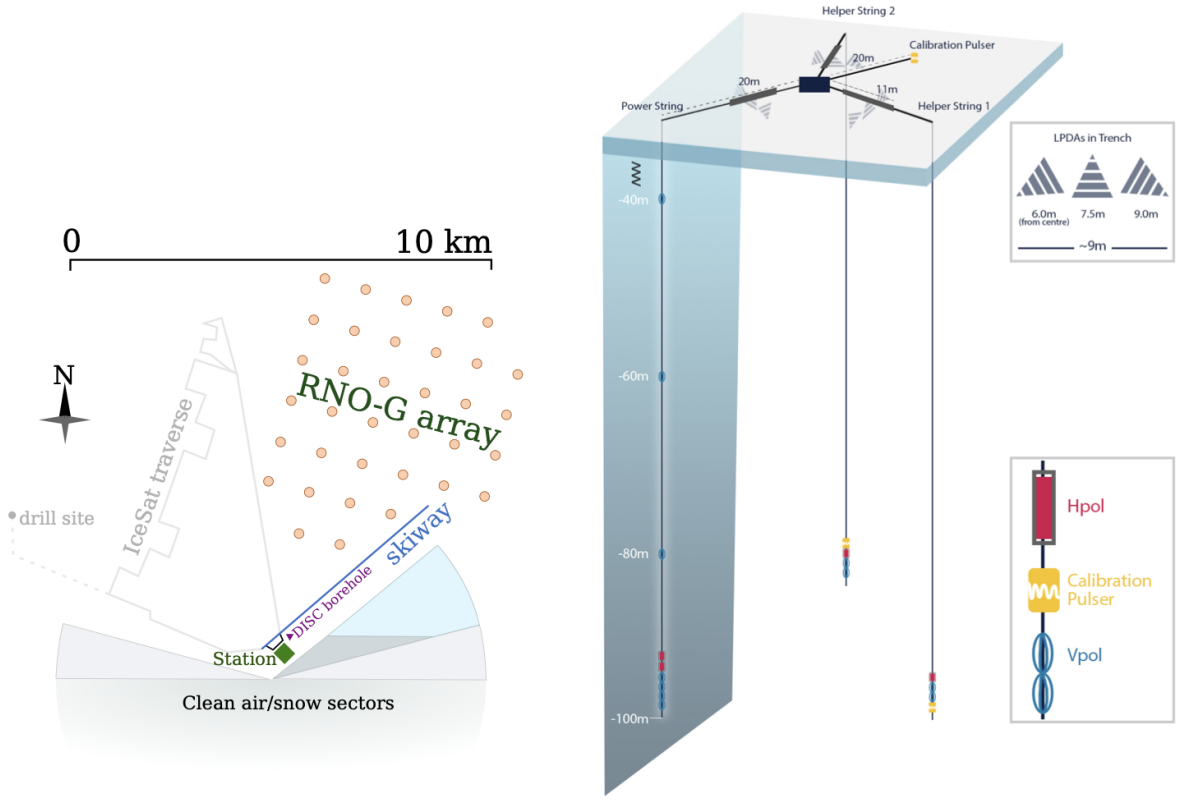


Fig. 2: Map of all planned stations of RNO-G together with schematic setup of a single station. Figure from Ref. [1].

The part of the station that is especially relevant for this thesis and in general for the detection of neutrinos is the deep component. This component consists of three strings (two helper strings and one power string) in the ice equipped with Vpol (vertically polarized) and Hpol (horizontally polarized) antennas and calibration pulsers. The heart of the detector is an interferometric phased array (in the power string), formed by four Vpol antennas spaced about 1 m apart. The phased array antennas are numbered sequentially, starting with the lowest antenna which is called channel zero. It is installed at a depth of about 100 m [1]. The phased array allows to reconstruct the angle θ at which the signal of the neutrino generated shower (respectively of other sources) reaches the detector through using the following simple, geometric formula [3]:

$$\cos(\theta) = \frac{\Delta \mathrm{t} \cdot \mathrm{v}}{\Delta \mathrm{z}} = \frac{\Delta \mathrm{t} \cdot \mathrm{c}}{\Delta \mathrm{z} \cdot \mathrm{n}} \tag{4}$$

where Δt is the time delay of the signal in different antennas of the phased array with vertical distance Δz. The velocity of the radio wave in the medium v is given by the speed of light in vacuum c, and n is the refractive index of the considered medium.

2.3 Weather Balloons and Radiosondes

The first unmanned balloon to reach altitudes above 12 km and still collect data such as temperature and barometric pressure was launched in 1893 by physicist Gustave Hermite [18]. Continuous improvements in both the balloon itself (for example, the use of rubber instead of paper) and the instruments used made it possible to study the structure of the atmosphere. At first, it was difficult to do research based on data collected by unmanned balloon flights because the instruments were hard to find after the flight. This changed after 1929 with the invention of radio transmitters that could send the collected data directly to a receiver. This type of weather balloon is now called a radiosonde. Since 1929, weather balloons have been an important tool for researchers studying the atmosphere and weather forecasting. Radiosondes have not been replaced by satellites because the sondes also collect data vertically, complementing the horizontal measurements of satellites [19]. In recent decades, weather balloon research has become increasingly important in the study of climate change [20].

The Integrated Characterization of Energy, Clouds, Atmospheric state, and Precipitation at Summit (ICECAPS) project explores the interaction between atmospheric processes and the Greenland Ice Sheet using different instruments. This includes atmospheric observations that are carried out by radiosonde carrying weather balloons operating twice a day [2]. A radiosonde is a device that measures for example temperature, atmospheric pressure and humidity and transmits the data using radio waves.

A modern ascending balloon launched from land or ship can reach altitudes of about 35 km before bursting due to barometric pressures of $3 - 5$ hPa. The instruments, i.e. radiosondes, carried by the balloon are safely returned to the surface by parachute [20]. Radiosondes can also be transported and dropped from airplanes without a balloon being needed. Again, a parachute is used to return the instruments to Earth. This type of radiosonde is called a dropsonde. There are also radiosondes that are carried by rockets [19]. The latter are not relevant for this work, but are mentioned for completeness. The balloon system currently used in the ICECAPS project consists of an ascending balloon carrying a radiosonde of type Vaisala RS41-SGP. This radiosonde continuously emits a radio wave signal [21]. A picture of the radiosonde and the balloon is shown in Figure 3.

(a) (b)

Fig. 3: Picture of (a) a balloon launched at Summit Station (ICECAPS project) and (b) corresponding radiosonde. The blue arrow points at the antenna of the radiosonde. [22]

In meteorology, there is a widely used communication system for radiosondes called the Meteorological Aids (MetAids) service. To use this service, the frequency of a radiosonde must be in the frequency bands assigned to MetAids. The two most important frequency bands are the 403 MHz and the 1680 MHz bands. Less common are frequencies around 1780 MHz. It depends on several factors, such as strength of upper winds, which frequency is chosen. For example, the 403 MHz band is preferred when strong upper winds are present often, as this band has propagation features that result in greater reliability over long distances [19].

In the following, the terms of radiosonde and balloon are used interchangeably.

2.4 Coordinate Systems

A question that has been around for centuries is how to correctly map the Earth and its surface. The simplest way is to use spherical coordinates, such as in the geographic coordinate system. In this system, a point is characterized by its longitude and latitude and ellipsoidal height. The longitude is an angular quantity that defines the east/west position of a point whereas the latitude indicates the position in north/south direction. The global positioning system (GPS) for example is based on the World Geodetic System of 1984 (WGS 84). For this system the equator is used as origin for the latitude and the meridian in Greenwich is used as reference for the longitude. Furthermore, stations with well defined coordinates on the Earth's surface are used additionally as reference points. In coordinate systems like this, using spherical coordinates, the calculation of distances can become quite complicated. For this kind of applications, Cartesian coordinates are favored. Therefore, a coordinate transformation is necessary. A system based on quasi Cartesian coordinates is the so called Universal Transverse Mercator (UTM) system. This system is based on a coordinate projection onto a cylindrical coordinate system. This system

is then divided into different zones having easting and northing as coordinates given in meters, which results in a Cartesian like coordinate system [23].

3 Zenith Angle Reconstruction Using a Single Phased Array

In this chapter basic methods and observations used in this thesis are explained in detail.

3.1 Radio-Frequency of the Radiosonde

As described in section 2.3, there are two typical frequency bands used in radiosonde operations. In Figure 4 it can be seen that the weather balloon flying above one of the RNO-G stations is emitting radio waves in the 403 MHz band. This is consistent with the radiosonde specification [2].

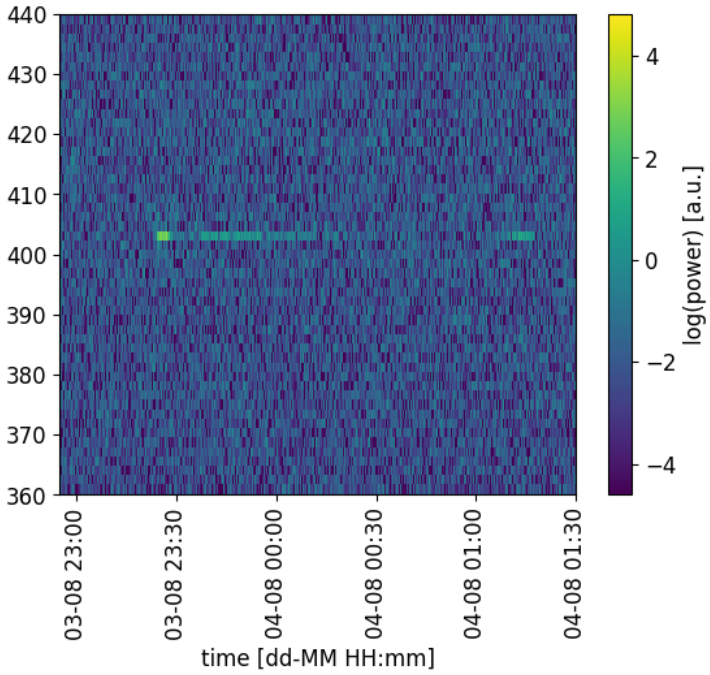
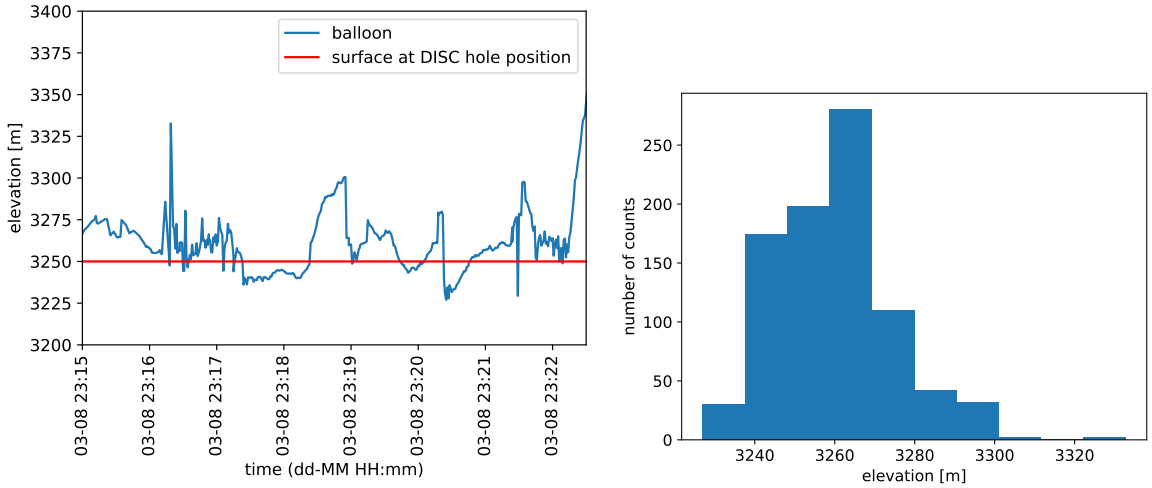


Fig. 4: Frequencies measured by station 23 over event times with natural logarithm of the power of signals as color bar

3.2 Specification of Surface Elevation

For some of the calculations in this thesis, the elevation (height above sea level) of the surface at the detector site is required. For example, to calculate the altitude (height above surface) of the weather balloon, since only its elevation is given in the GPS data. The surface at the Deep Ice Sheet Coring (DISC) hole has an elevation of about 3251 m [24][17], which will be used as the reference point in this thesis. To ensure that this corresponds to the elevation of the balloon, the balloon data before the start of the flight is to be checked. In Figure 3 the launch of a weather balloon at Summit Station is depicted. It can be seen that the balloon is launched from the height of a human elbow above the ground, which is in the range of about one meter above the surface. Therefore, the elevation in the data should be similar to the surface elevation. As shown in Figure 5 (a), the elevation before the launch varies in the order of tenths of meters.



(a) (b)

Fig. 5: Elevation of the balloon before the launch as time evolution together with elevation of the surface at the DISC hole (a) and as histogram (b).

For better illustration, the data points in this time interval are plotted in a histogram (cf. Figure 5 (b)). It can be seen that most of the values are in the bin of 3260 m. Since the accuracy of the height of the radiosonde used by ICECAPS is in the range of 10 m at best (cf. [21]), this is consistent with the elevation at the DISC hole. According to [21] the height of the radiosonde is calculated using its GPS coordinates. It is to be noted that the elevation value given by the radiosonde as well as its accuracy is actually given in geopotential meters (gpm). This unit corresponds furthermore to the so called geopotential height Z which is defined as:

$$\mathrm{Z} = \frac{1}{\mathrm{g}_0} \int_0^z \mathrm{g}\, dz'$$

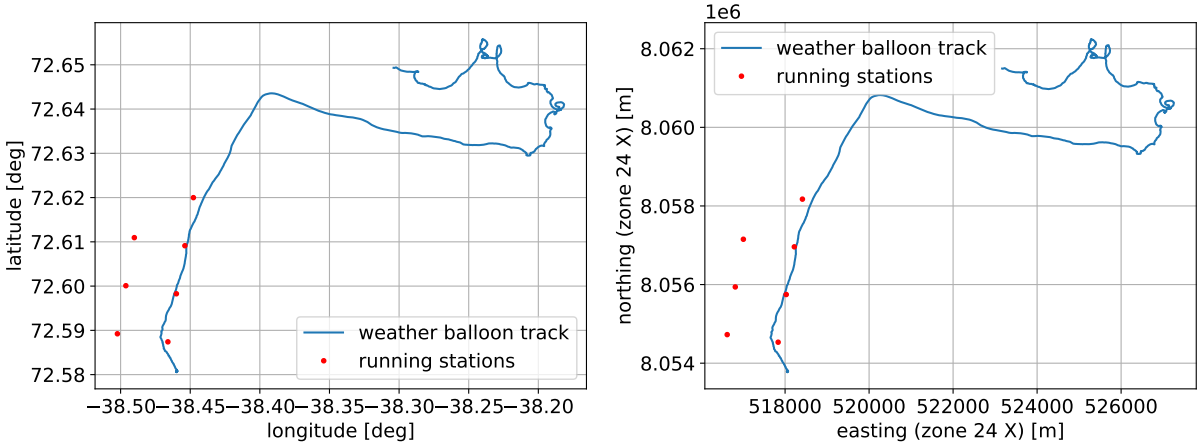
where z is the geometrical height, $\mathrm{g}_0 = 9.80665\,\mathrm{ms}^{-2}$ is the average acceleration of gravity at sea level and g is the local acceleration of gravity [25]. Therefore, the geopotential height Z equals the geometric height z when $\mathrm{g} = \mathrm{g}_0$ is given.

The surface elevation, which is given in geometric meters, is similar to the initial elevation of the balloon in geopotential meters. At this elevation, it can therefore be assumed that the difference between geometric meters and geopotential meters is expected to be negligible. The gravitational acceleration is decreased with increased altitude. Thus, the difference between geopotential meters and geometric meters is also increased with increased altitude. However, the change in the gravitational acceleration is small, e.g. for an increase of altitude of 100 km, the gravitational acceleration changes by $0.3\,\frac{\mathrm{m}}{\mathrm{s}^2}$, resulting in a deviation between geometric meters and geopotential meters in the order of four percent [26]. In this thesis, only altitudes of the balloon up to 3 km are considered. Since this value is much smaller than 100 km and thus the deviation between geopotential meters and geometric meters is much smaller than four percent, it will be neglected in the following.

3.3 Zenith Angle Calculation from Radiosonde Data

The balloon positions (GPS data) are given in the geographic coordinate system. As described in section 2.4, it is not useful to calculate distances in such coordinate systems. Therefore, it is necessary to transform the track data into a Cartesian system such as UTM. In this thesis, the python function `utm.from_latlon()` of the utm package is used. An example data set of the balloon positions in both coordinate systems can be seen in Figure 6. To get an impression

of the flight track in relation to the locations of the deployed detector stations, these are also shown in the plots. Each of these stations is associated with an index. The detectors on the left side are assigned the indices 11, 12 and 13 from bottom to top. The stations on the right side are labeled with indices 21, 22, 23 and 24, also from bottom to top. At the time of the balloon flight investigated in this thesis on 03.08.2022, station 12 was not taking data and is therefore not included in the subsequent analysis.



(a) (b)

Fig. 6: Example of balloon tracks in the geographic coordinate system (a) and UTM (b) including positions of the running stations.

To calculate the zenith angle between the stations and the balloon, it is necessary to determine the height difference h and the absolute distance d between stations and balloon. The position in the xy-plane and the elevation (in meters above sea level) of the balloon are given in the GPS data. The location of the stations is given in [27] and their elevation can be assumed as 3251 m (cf. section 3.2). Finally, the zenith angle θ can be trigonometrically calculated using:

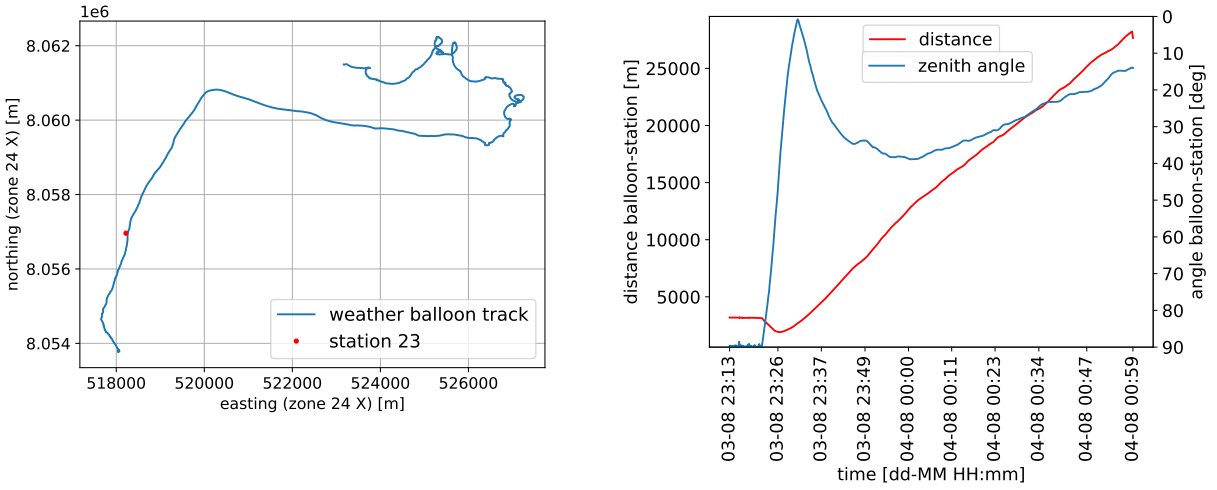
$$\cos(\theta) = \frac{\mathrm{h}}{\mathrm{d}} \tag{5}$$

An example of the time evolution of the zenith angle and the distance between the balloon and station 23 is shown in Figure 7 (b). For better illustration, the balloon track and the station in question are displayed in Figure 7 (a).

As described in section 2.2.3, the antennas are built in-ice. Therefore, the angle calculated by Equation 5 does not match the angle that is expected at the phased array. To estimate the angle at the phased array, the refraction between air and ice, as well as the density dependent refractive index of the ice itself (cf. section 2.2.2) have to be considered.

To handle this, the propagation module of the `NuRadioMC` python package is used [28]. The ice model "greenland_simple" with the properties mentioned in section 2.2.2 is already implemented within the package and is used for most parts of this thesis. To estimate the expected angle at the phased array, different surface interaction points of the radio signal emitted from the balloon are tested. For each interaction point x, the trigonometric zenith angle of the balloon in air θ_{tri} is given by:

$$\theta_{\mathrm{tri}} = \arctan\left(\frac{\mathrm{d} - \mathrm{x}}{\mathrm{h}}\right) \tag{6}$$



(a) (b)

Fig. 7: Example of balloon track over station 23 (a) and time dependence of zenith angle and distance between balloon and station 23 (b).

In this equation, d is the two dimensional distance between the balloon and the station (xy-plane) and h is the altitude (height above surface) of the balloon. Then, a ray tracing is performed between the surface interaction point on the surface and the position of the phased array. The depth is assumed to be $-98\,\mathrm{m}$, which is the depth of the center of the phased array. The best surface interaction point is given by the smallest squared difference between the refracted trigonometric angle in the ice calculated by Snell's law (Equation 3) of θ_{tri} and the launching angle of the ray tracing solution. This approach is similar to a χ^2 minimization with the difference that in this case there are no weighting factors included which are part of a usual χ^2 function (cf. section 5.3.1). The receiving angle of the respective ray tracing solution is finally the angle at the phased array. An example of this calculation compared to the trigonometric determination of the angle between station and balloon without ice (cf. Equation 5) can be seen in Figure 8. This shows that the expectation of the angle at the phased array indeed deviates from the basic trigonometric calculation without taking the ice into account.

A simpler approach to calculate the expected angle at the phased array would be to just take the angle given by applying Snell's law to the angle calculated in Equation 5 with the refractive index of the ice at the depth of the phased array. This is justified by the idea that the wave emitted by the balloon (approximated as a plane wave) is refracted by infinitesimal layers in the ice with infinitesimal differences in refractive index until the wave reaches the detector. This results in the total refractive index being equal to the refractive index at the position of the phased array.

In Figure 9 the results for both calculation methods are shown. For the trigonometric approach, the results are displayed for different refractive indices. It can be seen that the angles obtained by the ray tracing technique are quite similar to those determined trigonometrically together with Snell's law for a refractive index between the values $n = 1.70$ and $n = 1.75$. This result is consistent with the fact that the refractive index of ice at a depth of approximately $-98\,\mathrm{m}$ (depth of the middle of the phased array) in the ice model "greenland_simple" (cf. Equation 2) is about $n = 1.743$ [28]. Since the ray tracing method takes much longer to compute than the trigonometric method, but results in a similar angular development, the trigonometric calculation is favored. In the following, a refractive index of $n = 1.74$ will be used for the calculation of the zenith angle from the balloon data, unless otherwise specified in the text.

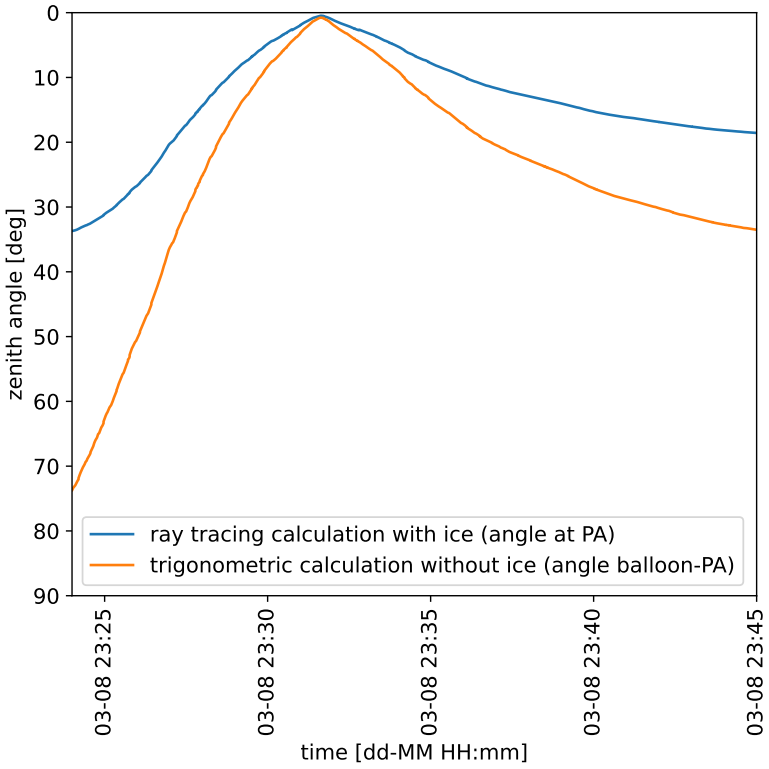


Fig. 8: Example of trigonometric angle calculation without ice and ray tracing calculation with ice.

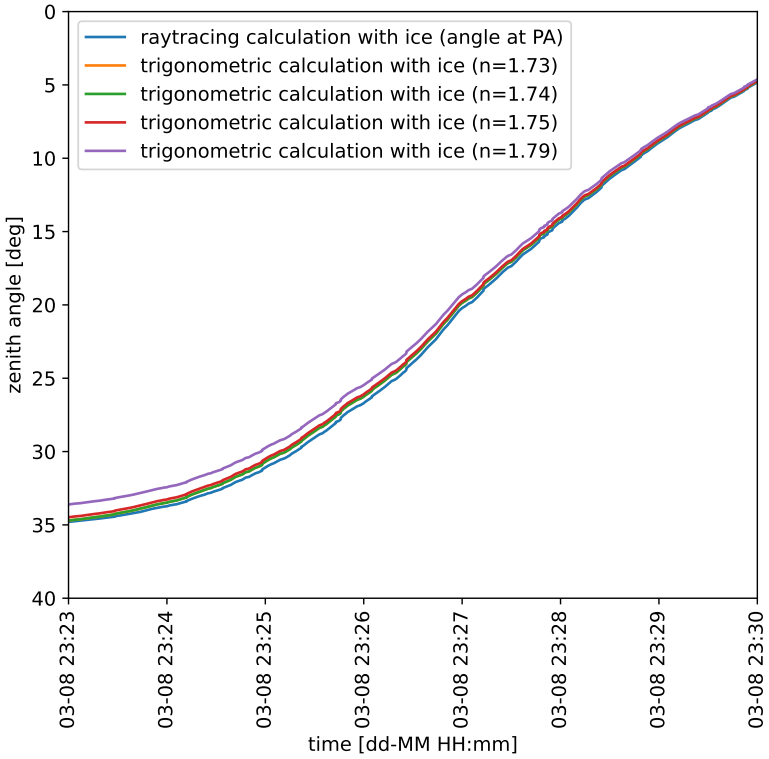


Fig. 9: Comparison of expected angles at the phased array for different calculation methods.

3.4 Distance, Angle and Power Development During Balloon Flight

In Figure 10, the distance between the balloon and the station, the corresponding angle and the power in the 403 MHz band measured by channel 0 are depicted in dependence on the time. It is visible, that the power values in the 403 MHz band are near 1 a.u. before the balloon starts collecting GPS data. This can therefore be interpreted as background noise. In channel 0, the power shows a peak when the distance between the balloon and the station is minimal and the corresponding elevation, i.e. the angle above the horizon, is near maximum, in contrast to the zenith angle shown in the plot. The zenith angle is zero when the balloon is directly above the station.

Furthermore, it is apparent that the power is decreased to the level of noise background when a zenith angle of zero degrees is reached. A reason for this behavior is discussed in section 3.5.1. After this minimum, the measured power is then increased again for decreased zenith angles, but it should be noted that the signal-to-noise ratio (SNR) is significantly lower than for the power peak at 23:25.

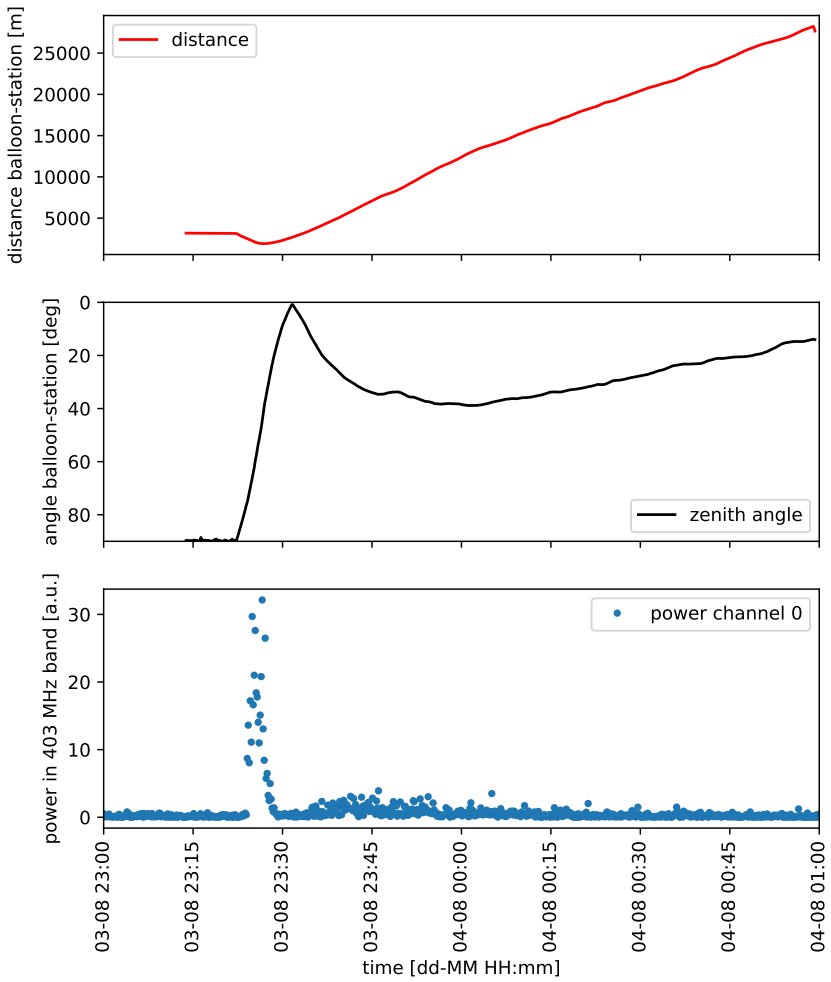


Fig. 10: Example of the angle and distance between the balloon and the station, and the power of the phased array channel 0 over the same time period.

3.5 Angle Determination from Phased Array Data

In this section, the general procedure for the angle determination from the phased array data is explained. Furthermore, a method to examine the calibration of the cable delays and the spacing of the phased array antennas is described.

3.5.1 Angle Determination using Delay-and-Sum Beamforming

The phased array is the important part of the RNO-G stations used in this thesis, as mentioned in section 2.2.3. To extract the angle of the signal from the detector data, a similar method is used as for the so-called delay-and-sum beamforming trigger described in [3].

The general delay-and-sum beamforming technique is based on calculating a coherent sum $\mathrm{S}(t)$ of the signals of an antenna array. This sum can be expressed by [3]:

$$\mathrm{S}(t) = \sum_{\mathrm{n}=0}^{\mathrm{N}-1} \mathrm{w_n y_n}(t - \delta_\mathrm{n})$$

where N is the number of antennas, $\mathrm{w_n}$ is a weighting factor for the antenna amplitude, $\mathrm{y_n}$ is the time-dependent antenna signal which will be referred to as trace in the following and δ_n is the delay that is applied to the respective signal.

Using the same weighting factors $\mathrm{w_n}$ for all antennas has the advantage that the uncorrelated noise, e.g. the thermal noise background, only scales with $\sqrt{\mathrm{N}}$, whereas the correlated signal is included in the calculation with a scaling of N [3]. In reality, using the same weighting factors is justified in most cases, as the same hardware is used for the channels.

Since the phased array used by RNO-G consists of four antennas spaced one meter apart, the incoming signal has a geometric delay ($\Delta \mathrm{t_{geo}}$) between the respective antennas. Based on Equation 4, this delay can be calculated using

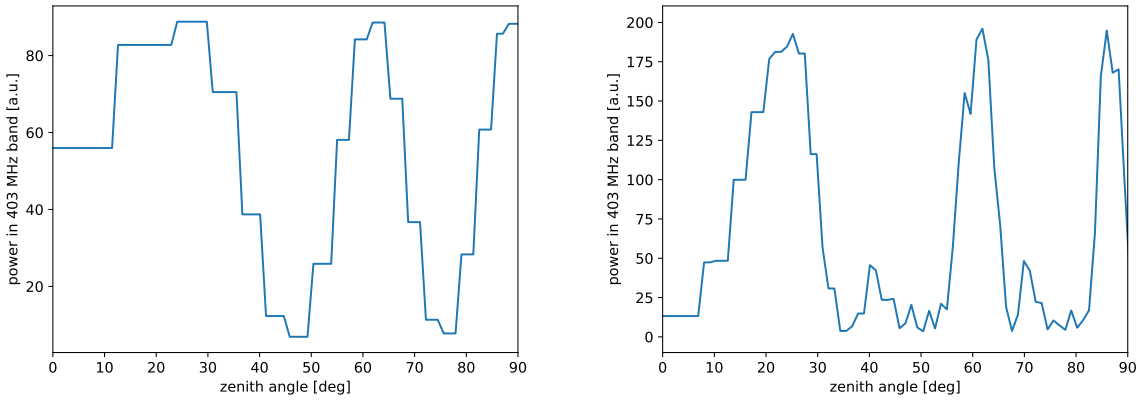
$$\Delta \mathrm{t_{geo}} = \frac{\mathrm{z_{max}} - \mathrm{z_{antenna}}}{\mathrm{c}} \cdot \mathrm{n} \cdot \cos(\alpha) \tag{7}$$

where $\mathrm{z_{antenna}}$ is the z-position of the considered antenna and $\mathrm{z_{max}}$ is the z-position of the highest phased array antenna. In the case of RNO-G stations, this is the z-position of the channel 3 antenna. Furthermore, c is the speed of light, n is the refractive index of the medium (ice) at the depth of the phased array. The latter is assumed to be constant at the depths of all phased array antennas. Finally, α represents the angle of the incoming signal. This angle is unknown and will be determined by the following process which later on will be referred to as phasing.

The geometric delay is not the only delay to consider. Due to the different lengths of the antenna cables and therefore different propagation times of the signals, a delay caused by the cable must also be considered in the calculations. The total time delay between the antennas is therefore given by $\Delta \mathrm{t_{total}} = \Delta \mathrm{t_{geo}} + \Delta \mathrm{t_{cable}}$. With this, a time delay for each channel can be calculated in dependence of the angle of the incoming signal.

To determine the optimal angle, the time dependent signals measured by the channels are shifted by the calculated delays. Since the `numpy.roll()` function in python moves elements to the front that are rolled behind the last possible element, at least the number of samples corresponding to the total delay have to be cut off. After cutting, the traces are summed and the power of the spectrum is calculated. Since in this project only the area around 403 MHz is relevant, the power is calculated only in the 403 MHz band. Finally, the powers are compared for different delays and therefore for different angles of the incoming signal. The optimal estimated angle is theoretically given by the highest power in the 403 MHz band. In reality, there are often multiple angles that result in similarly high powers, which does not lead to an unambiguous angle reconstruction. An example of this is shown in Figure 11, where the calculated power for different time delays and thus different angles is shown as a function of angle for one event. In both cases, for two and for four channels, side maxima can be found. It should be noted that only zenith angles between 0 and 90 degrees are considered, as the balloon angle is known to

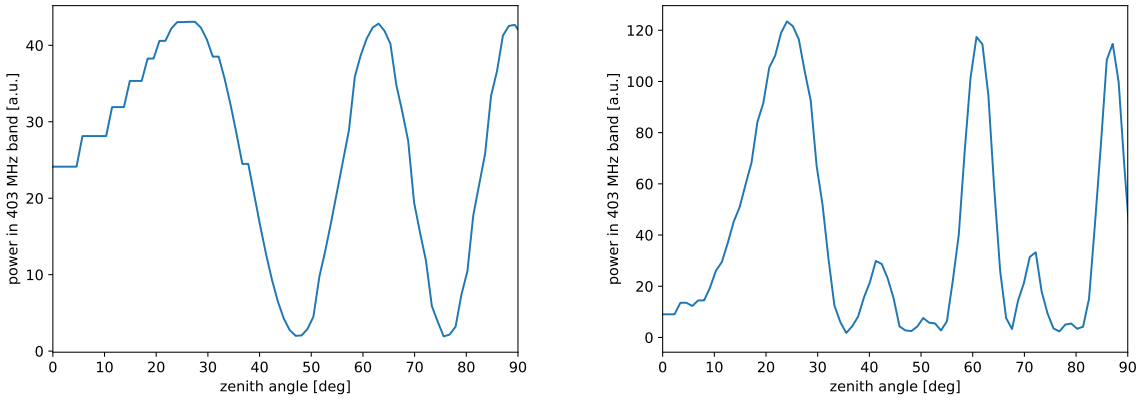
be within this range. A further limitation of the angles to 0 to 35 degrees is possible, as larger angles at the phased array are unphysical for a balloon flying outside the ice (cf. section 2.2.2), and will be considered later on. However, in order to detect periodic behavior in the angle reconstruction, angles in the range of 0 to 90 degrees are considered for the moment.



(a) (b)

Fig. 11: Example of power in 403 MHz band over different possible angles (a) for two channels, (b) for four channels.

This process can be optimized by upsampling in frequency- space (increasing the sampling rate) and filtering the traces, i.e. the time dependent signal measured by the channels. Upsampling results in a more precise shifting of the signals, since the trace can only be rolled by integer samples using `numpy.roll()`. Applying a bandpass filter in the range between 350 MHz and 450 MHz reduces the noise in the 403 MHz band when summing the traces. The filtered and upsampled version of the plot in Figure 11 is shown in Figure 12. As can be seen, the peaks are much smoother and have less power compared to the peaks in Figure 11. The lower power is a result of the filtering and upsampling process. The filtering reduces the noise whereas the upsampling leads to a narrower bin at 403 MHz. Both lead to a decrease in the power of the peaks.



(a) (b)

Fig. 12: Example of power in 403 MHz line over different possible angles (a) for two channels, (b) for four channels after filtering and upsampling.

In the following step, all zenith angles between 0 and 90 degrees are plotted over events occurring

in the interval between 23:20 and 23:45, with a color bar encoding the power in the 403 MHz band. This allows to visualize the angle evolution over time (event times), so that it can be compared to the expectation calculated as described in section 3.3 (cf. Figure 13). As can be seen in Figure 13, the evolution of the phased data matches the evolution of the expected angles in the time range near 23:25 and near 23:40. At other times, e.g. between 23:30 and 23:35, no shifting of the signals leads to an significant increase in the power in the 403 MHz band and therefore the angle of the signal cannot be estimated. This is not due to inaccuracies in the phasing procedure, but to the measured power values in the antennas as can be seen in Figure 10.

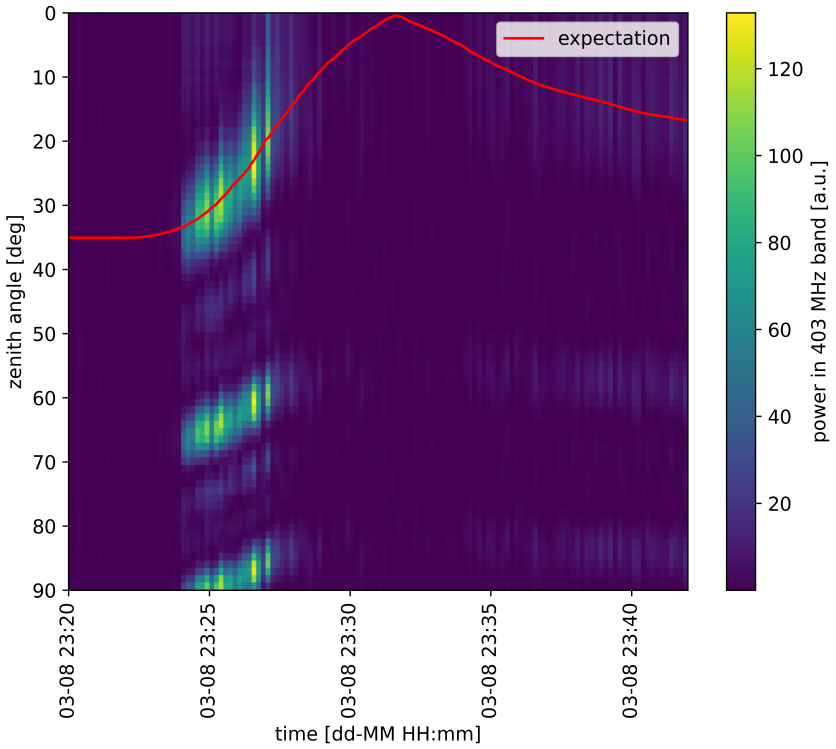


Fig. 13: Time evolution of zenith angles with trigonometric expectation of the zenith angle in-ice.

From Figure 7 follows that the balloon starts collecting GPS data at 23:13. The actual flight begins at about 23:22. The first clear signal in station 23 is observed at approximately 23:24 (cf. Figure 13). This discrepancy between the visible signal and the time of the balloon launch cannot be explained by the distance between the balloon and the station and the corresponding decrease in signal amplitude as a signal exceeding the noise is measured at 23:45, when the distance is larger than at 23:22 (cf. Figure 10). This should be investigated in more detail. However, it is beyond the scope of this thesis.

For expected zenith angles of zero degrees, i.e. when the balloon is directly above the station, the angle resulting from the phasing process is not unique due to an overall low power in the 403 MHz band. An explanation for this can be found by looking at the vector effective length (VEL) of the phased array antennas as a function of the zenith angle (cf. Figure 14). The VEL is a measure for the direction dependence of the antenna sensitivity as it connects the incoming electric field to a respective response voltage [29]. From Figure 14 it can be seen that the absolute value of the VEL and thus the sensitivity of the antenna to a radio signal received at zenith angles of about 0 and 180 degrees is minimal. Therefore, a strong signal at times when the balloon is directly above the station is not expected and can explain the absence of a signal in this range.

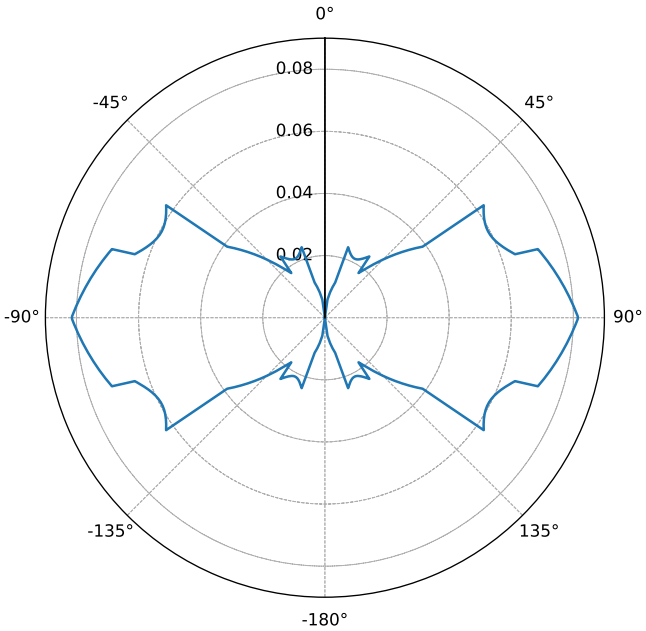
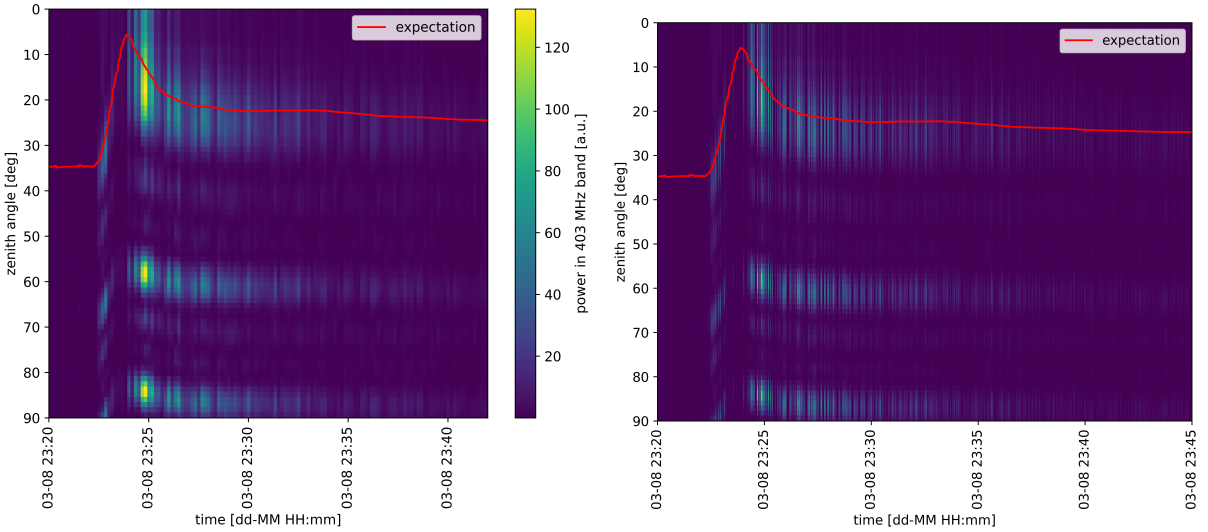


Fig. 14: Vector effective length of the phased array antennas at 403 MHz for an antenna model of the Vpols deployed in RNO-G over zenith angle.

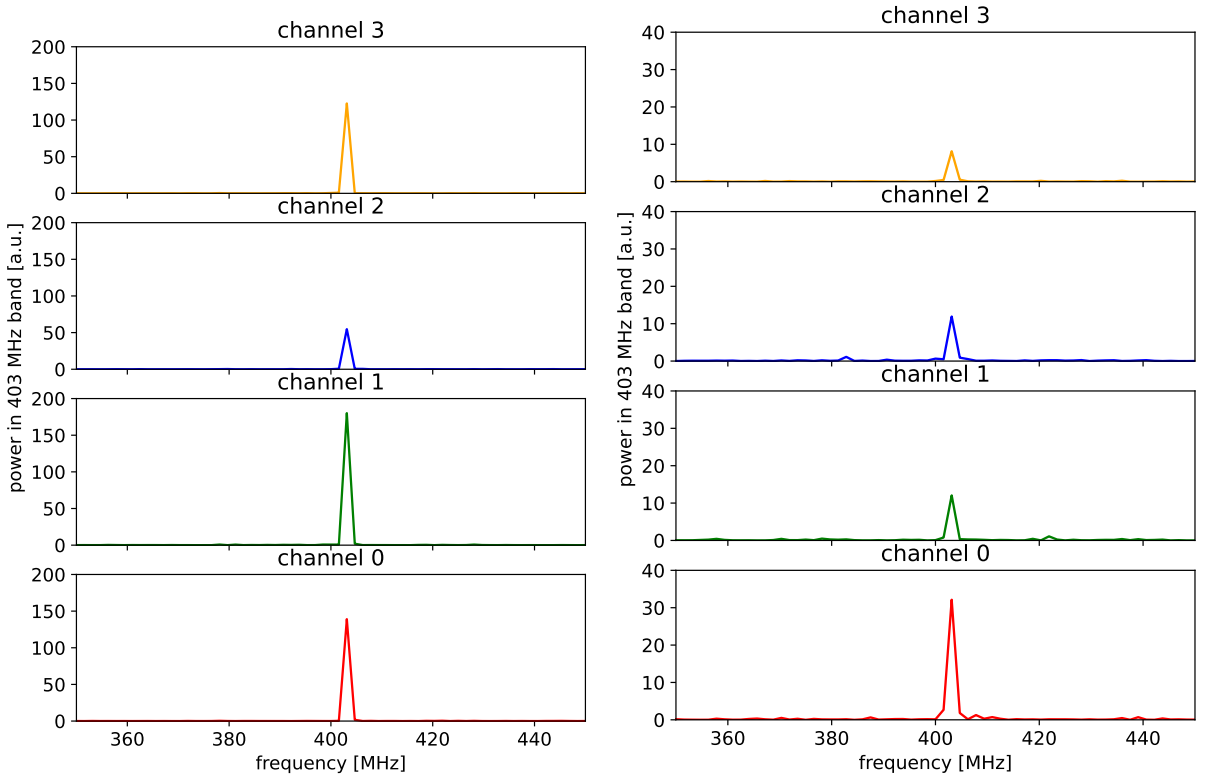
In addition, the radiosonde most likely also has a dipole-like emission pattern [19]. This contributes also to the absence of a strong signal when the balloon is directly above a station. It should be noted, however, that the radiosonde is not always perpendicular to the ground due to air flow in the atmosphere causing the radiosonde to undergo deflection.

The data used so far was only data generated by forced triggers with a rate of 0.1 Hz [1]. To test whether the missing angles can be reconstructed using more data points, the phasing process can be applied to the data generated by any triggered event. In Figure 15 it can be seen that using the data from all triggers does not result in more information about the angles in the critical time intervals. It should be noted, that in Figure 13 the data of station 23 was examined. In Figure 15 station 21 is considered. For this station the observed angles also match the expectation.
When looking at the peak power values of station 21 and 23 it is apparent that they are similar (cf. Figure 15 (a) and Figure 13). However, at station 11, for example, the maximum power after the phasing process exceeds 1000 a.u., which is approximately a factor of 100 larger than the values at stations 21 and 23. To rule out that this difference arises as a result of inconsistencies, e.g. incorrect rolling during the phasing, the power values in the different channels of the stations are checked. The power in different channels of station 11 and 23 of events leading to the highest observed power are depicted in Figure 16. The initial power values on which the phasing is performed defer from one another. The values for station 11 (Figure 16 (a)) are in a range between 50 and 170 a.u., while the power of the individual channels in station 23 only reach values up to 30 a.u.. Therefore, the power observed after phasing is not a measure of the quality of the process. The difference in power is more likely to be an effect of varying antenna sensitivities in dependence of the zenith angle and the variable emission pattern of the radiosonde emitter (possible deflection of the sonde due to wind).



(a) (b)

Fig. 15: Time evolution of zenith angles with trigonometric expectation of the angle in-ice (a) only forced triggered data, (b) data from any trigger



(a) (b)

Fig. 16: Power measured in different channels over frequency for events generating the highest power (a) in station 11, (b) in station 23.

For the stations investigated so far, the observed angles after the phasing are in good agreement with the expected values. The time evolution of the angles for station 13 again with the power in the 403 MHz indicated as color is depicted in Figure 17. It can be seen that the expected angles are larger than the angles with the highest power observed after the coherent summation of the antenna signals. A possible reason for this is discussed in section 3.5.2. Moreover, it is apparent that a reasonable signal is only observed at a few times in this station. This may be due to an generally lower sensitivity of the antennas as the maximum power only reaches values up to 50 a.u..

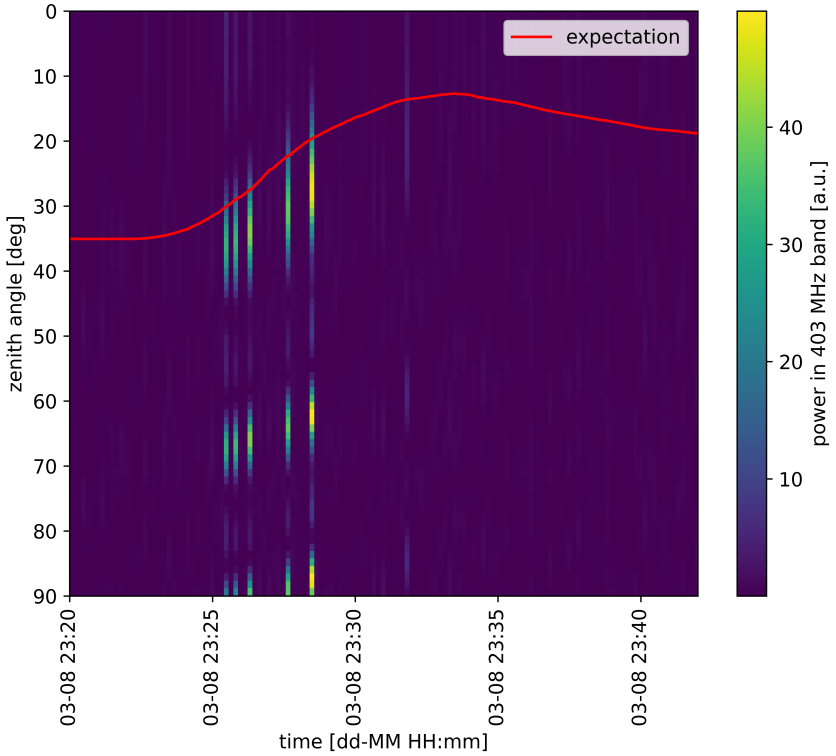


Fig. 17: Time evolution of zenith angles with trigonometric expectation in-ice for station 13.

The corresponding plots of the time evolution of the zenith angles for stations 11, 22 and 24 can be found in Appendix A.

3.5.2 Calibration Study using the Beamforming Technique

The beamforming technique can be applied to various combinations of two channels at a time. Incorrect cable delays can be identified if the phasing angle is always showing an offset when a particular channel is included in the calculation. Additionally, a regular spacing between the antennas can be verified. For regularly spaced antennas, it is expected that the number of possible angles will increase as the distance between two particular channels is increased, since the number of wavelengths matching the distance is also increased.

In Figure 18, the power in the 403 MHz band after phasing of different channel combinations of station 23 is depicted in dependence of the possible zenith angles. The number of complete peaks for phasing between channel 0 and 1 is the same as for phasing between channels 1 and 2 and between channels 2 and 3. The wavelength of the radio signal emitted by the radiosonde is approximately given as $\lambda = 0.4\,\mathrm{m}$. At the time of interest, the angle of the incoming signal is approximately 30 deg. With an assumed distance of 1 m between the antennas, the distance for the incident wave becomes about 1.15 m. Therefore, more than two wavelengths fit between the antennas and thus two complete peaks and the beginning of a third peak are expected. This

matches the observed number of peaks as per Figure 18. Additionally, phasing between channels 0 and 2, respectively between channels 1 and 3 result in twice as many peaks as the phasing between channels spaced only one meter apart. This matches the assumption of the antennas being equally spaced.

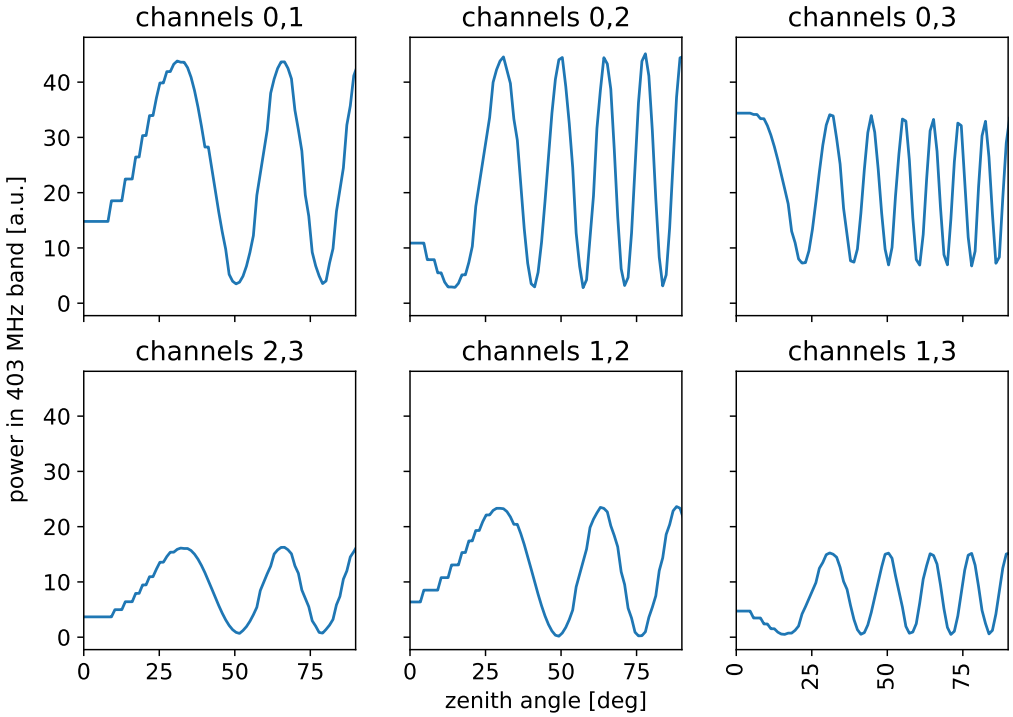


Fig. 18: Power after phasing in 403 MHz band over zenith angle for different channel combinations.

The power in the 403 MHz band after applying the phasing technique in dependence of the zenith angle for the same channel combinations as Figure 18 is shown in Figure 19. In addition, the power after coherently summing the signals of all four channels is also depicted (cf. Figure 19). It can be seen that there is no channel that causes an obvious shift in the peak positions as the power of all channel combinations peaks at the same angle. Excluded from this statement are the peaks that only occur at larger distances between the antennas, since no peaks are to be expected at these positions at smaller distances. This indicates that the values of the cable delays used for the phasing are reasonable for station 23.

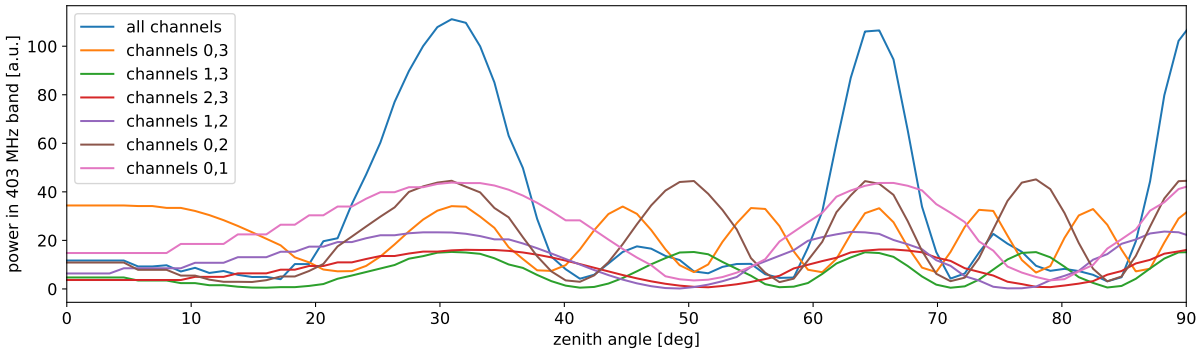


Fig. 19: Power after phasing in 403 MHz band over zenith angle for different channel combinations including phasing for all channels in station 23.

As described in section 3.5.1, the observed angles in station 23 match the expectation (cf. Figure 13), whereas the angles with the highest power values in station 13 deviate from the expected angles (cf. Figure 17). To determine if this deviation is induced by incorrect cable delays, the power after phasing for different channel combinations of station 13 is shown in Figure 20. A line indicating the expected angle at the event time is also plotted. It can be seen that the maxima of the phasing peaks of different channel combinations are shifted with respect to each other. Furthermore, it is noticeable that the phasing of two channels each results in a higher deviation from the expected angle, i.e. the maxima of the peaks are shifted towards higher angles. This may be caused by incorrect cable delays. To verify this, an investigation of several balloon flights is necessary to ensure reproducibility of the issue.

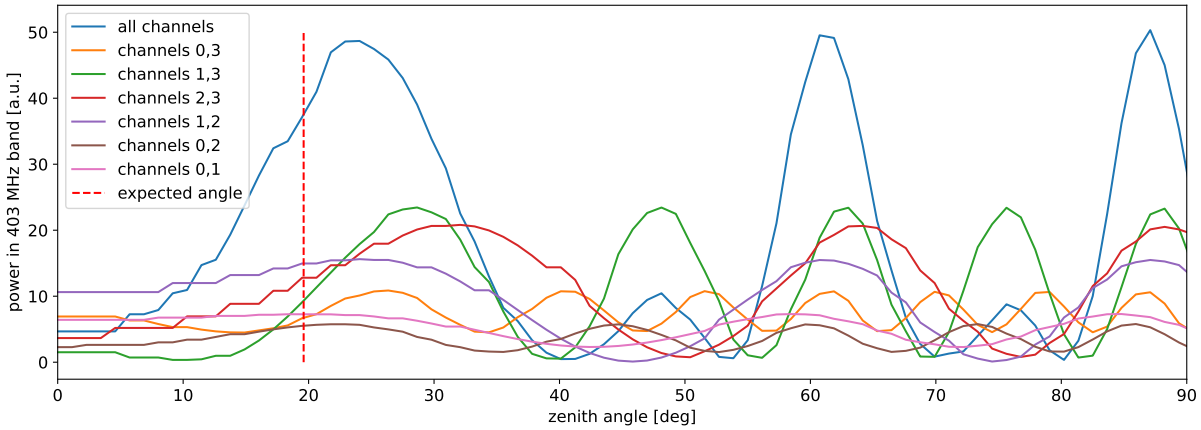


Fig. 20: Power after phasing in 403 MHz band over zenith angle for different channel combinations including phasing for all channels in station 13. Indicated is also the expected angle at the event time as dashed red vertical line.

The phasing of different channel combinations can be used to quantify potential deviations in the reconstructed angles at the phased array caused by single antennas. For this purpose, two channel combinations (e.g. channels 0,1 and channels 1,2) and several times can be selected. For each chosen time, the mean angle of the power peak with the highest power value can be inserted in a diagram. For an ideal calibration it is expected that all data points are positioned on a straight line with a slope of 1. If this slope deviates from 1, inconsistencies in the spacing of the antennas and the cable delays are expected, since the angles obtained at the phased array depend on these quantities (cf. Equation 4). Repeating the procedure for other channel combinations allows to identify the channel producing incorrect angles and to estimate the significance of the inaccuracy (deviation of the slope from 1). This was not done explicitly, as it would have been beyond the scope of this thesis.

4 Calibration Study Using the Refractive Index

As described in section 3.5.1, the refractive index of ice is required to calculate the time delays of the signal between the phased array antennas. The most probable phasing angle (angle with the highest power in the 403 MHz band) can now be estimated for different refractive indices and compared to the expected angle given by applying Snells law with the same refractive index to the trigonometric calculation of the balloon position (see section 3.3). The refractive index with the smallest difference between the calculation and the measurement can be stored. Finally, the collected refractive indices for several times during a balloon flight can be displayed in a histogram. In Figure 21 an example is shown for station 23. The bin in the histogram with the largest number of counts, respectively the mean value of a Gaussian distribution fitted to the histogram data, is expected to correspond to the real value of the refractive index of the ice at the depth of the phased array for an ideal calibration. I.e. if the index of refraction estimated

from the histogram deviates significantly from the known index of refraction at the depth of the phased array, an issue regarding the antenna spacing should be suspected.

In Figure 21 it can be seen that the refractive index with the most entries is $n = 1.73$. Additionally, the Gaussian function fitted to the data has a mean value of 1.730 and a standard deviation of $\sigma = 0.010$. This matches the expected refractive index of $n \approx 1.74$ provided by the ice model 'greenland_simple' at the depth of the phased array. It should be noted that a number of 24 data points is not sufficient to make a conclusive statement, but the analysis of more data is necessary (taking into account different balloon flights). The analysis for station 22 results in a refractive index of $n_{22} = 1.747 \pm 0.014$ (15 data points). The uncertainty is given by the standard deviation and is therefore only statistical. This is in good agreement with the expectation. However, for stations 11, 13, and 21 the result deviates more from the expectation as this analysis yields values of $n_{11} = 1.718 \pm 0.011$, $n_{13} = 1.72 \pm 0.04$ (5 data points) and $n_{21} = 1.708 \pm 0.007$ (31 data points). The number of data points considered varies from station to station. This is because stations show reasonable signals at different numbers of time points. Station 24, for example, was excluded from this analysis because it only shows a signal at three times.

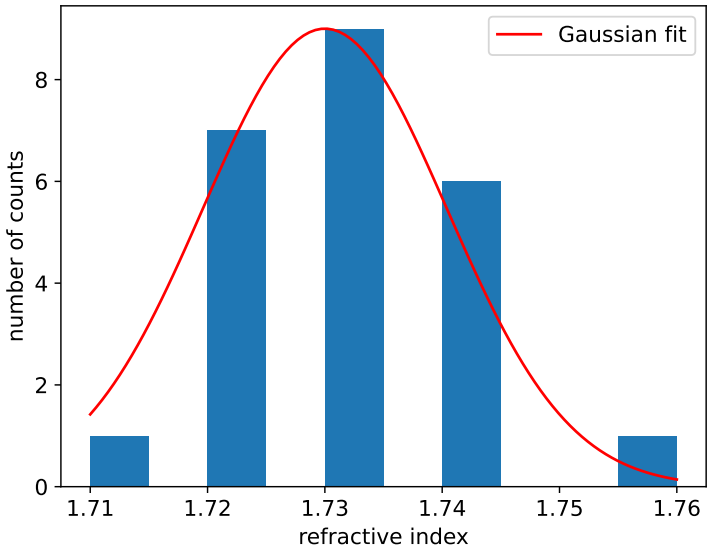


Fig. 21: Histogram of refractive indices with smallest deviation between expectation and measurement with Gaussian fit for station 23.

With more data, this offers an option to detect inconsistencies in the antenna spacing in different stations. However, given the results based on the few data points described here, it may be useful to carefully re-examine the distances between the phased array antennas of stations 11, 13, and 21.

5 Radiosonde Tracking

First in this section, two-dimensional considerations for balloon trajectory reconstruction are explained. Then, a method for a three-dimensional track reconstruction is described. In addition, an error calculation is demonstrated and the overall results of the tracking are discussed.

5.1 Phased Array Angle to Radiosonde Position in the xy-Plane

Before investigating the full three-dimensional case, the simpler two-dimensional case is considered in a first step. The goal of this investigation is to determine if the position of the balloon

can be reconstructed for a given balloon elevation.

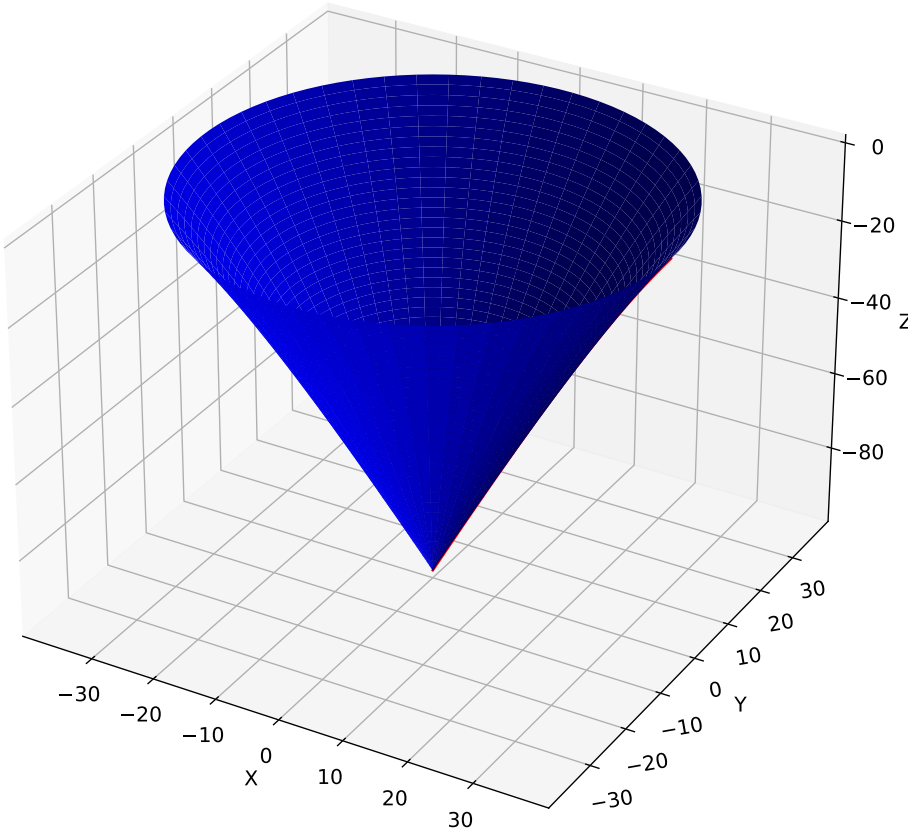


Fig. 22: Sketch of a curved cone above the phased array due to nonlinearity of ice.

In an ideal case, the coherent summation of the signal of the different channel antennas as described in section 3.5.1 results in a well-defined angle at the phased array. The source of a wave impinging at this angle is located on a cone surface. The opening angle of the cone corresponds to the measured angle at the phased array, assuming a constant index of refraction. When multiple stations are considered, the balloon is expected to be located at an intersection of the different cones.

The assumption of a constant refractive index is not justified in reality (cf section 2.2.2). Therefore, further steps are required to estimate the position of the balloon. First, the nonlinearity of the index of refraction in ice must be taken into account. The part of the cone which is located in the ice therefore has a curved surface. This curvature can be determined by performing a ray tracing between the phased array and possible surface interaction points. The best surface interaction point, and thus the curvature, is given by the smallest squared difference between the starting angle of the ray tracing and the angle given by the phasing procedure. An example of such a cone can be seen in Figure 22. It must be noted that the coordinate system is shifted in such way that the station is positioned at the origin of the xy-plane.

Besides the nonlinearity of ice, the transition from ice to air must also be considered. In this calculation the wave is assumed to propagate in the opposite direction as in reality. For the transition, Snell's law (Equation 3) must be applied on the receive angle of the ray tracing performed to get the surface interaction point. In air, a truncated cone must be considered.

The radius r_{top} of the top circular base can be calculated by:

$$r_{top} = \rho + h \cdot \tan(\theta) \tag{8}$$

with ρ describing the radius of the bottom circular base of the truncated cone in air, the altitude of the balloon h at the given time and the receive angle θ of the ray tracing in air. An example of the realistic version of a composed cone including both the nonlinearity in ice and the transition to air is shown in Figure 23.

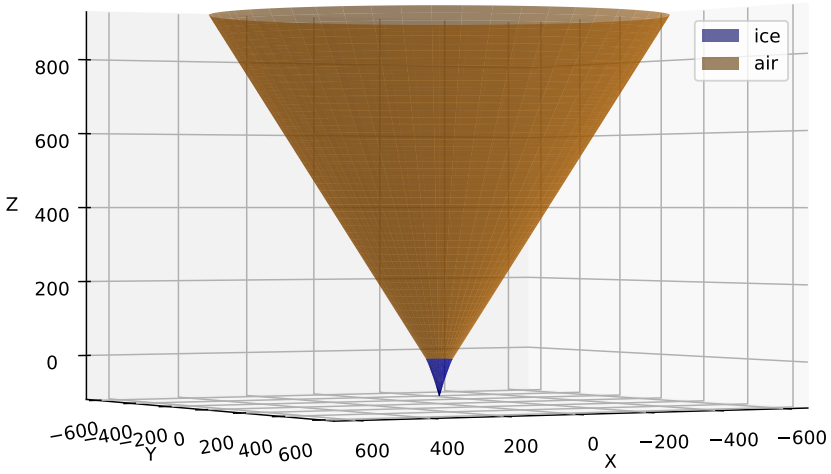


Fig. 23: Sketch of a truncated cone in air on top of curved cone in ice above the phased array.

As mentioned earlier, in reality it is often not possible to estimate a well-defined angle at the phased array. This can be seen again by looking at the power distribution over the possible zenith angle at the event of interest. The peaks with the highest powers in Figure 24 have a certain width. Through fitting a Gaussian function to the observed peaks, the mean value of the peak m as well as the width of the peak σ can be estimated. The Gaussian function is given by:

$$f(x) = a \cdot \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) + d \tag{9}$$

with additional fitting variables a (amplitude) and d (displacement in y direction).

To investigate whether the top circular bases of the cones intersect at a given time, the fitted variables from the Gaussian fit can be used to calculate the radii of the corresponding circles at the altitude of the radiosonde at that time. In this analysis, the one-standard-deviation uncertainties σ of the mean angle are taken into account. This transforms a circle into a circular ring.

The angle at the phased array corresponding to the inner radius of the circular ring is given by $\alpha_{\mathrm{i,inner}} = \mathrm{m_i} - \sigma_{\mathrm{i,inner}}$. Analogously, the angle at the phased array corresponding to the outer radius of a circular ring is $\alpha_{\mathrm{i,outer}} = \mathrm{m_i} + \sigma_{\mathrm{i,inner}}$, where i denotes the different stations. The relevant values of the fitted parameters are listed in Table 1. A complete list of all Gaussian fit parameters can be found in Appendix B.

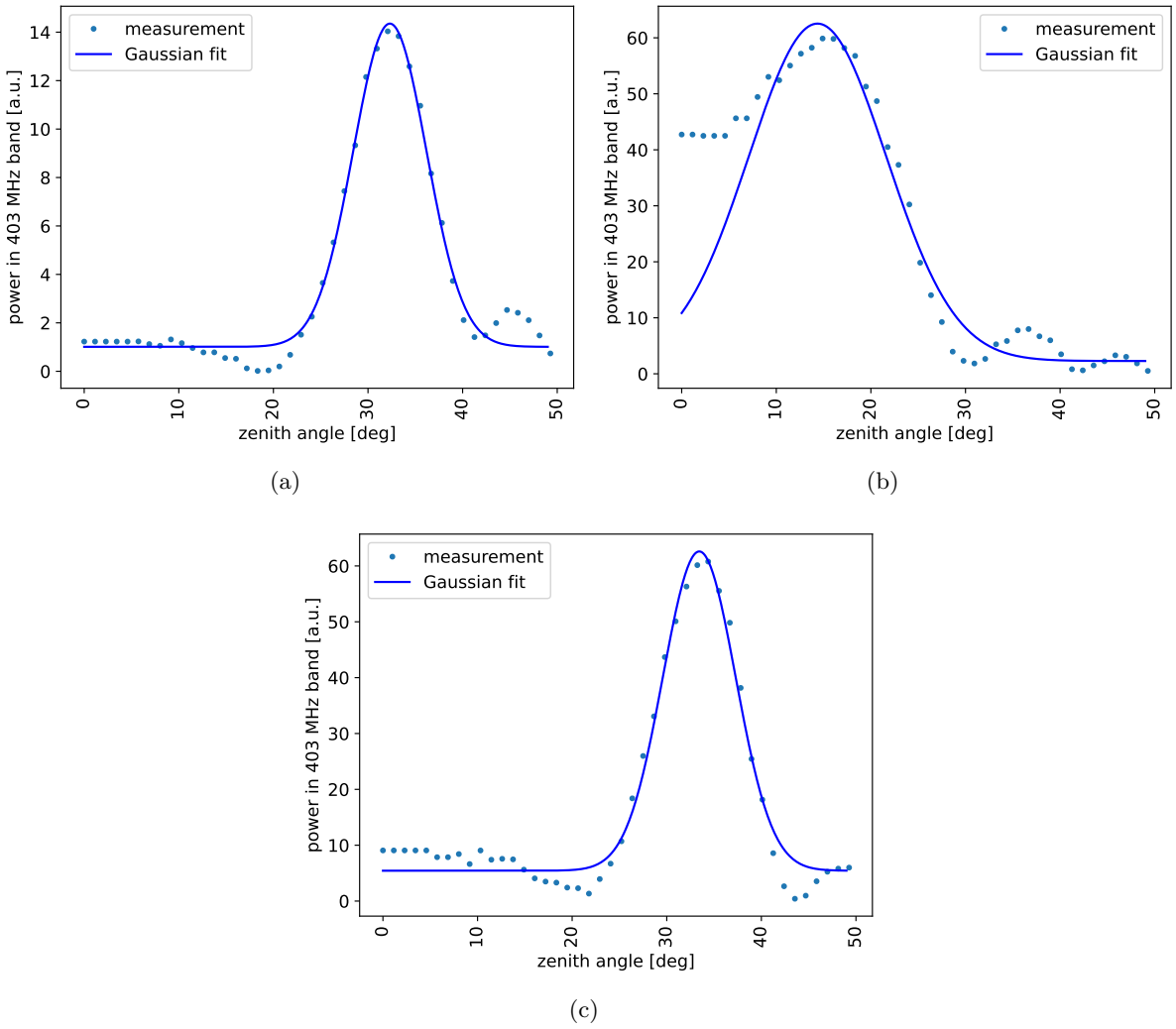


Fig. 24: Power over possible zenith angles at the phased array (phasing) at the same time for station 11 (a), station 21 (b) and station 23 (c). For all stations a Gaussian fit to the peaks is also plotted.

Tab. 1: Relevant fitted values of Gaussian functions for stations 11, 21 and 23

Station number	m [deg]	Δm [deg]	σ [deg]	$\Delta\sigma$ [deg]
11	32.31	0.10	3.87	0.11
21	14.355	0.26	7.3	0.3
23	33.43	0.24	3.85	0.30

The circular rings for stations 11, 21 and 23 are investigated at 23:24. The times of the data points taken by the stations deviate up to six seconds from each other. A measurement is triggered every ten seconds in all stations (cf. section 3.5.1). Therefore, data points measured within an interval of ten seconds are considered as taken at the same time.

For stations 11 and 23 in this example, the value for $\alpha_{\text{i,outer}}$ would exceed 35 deg. Such an angle cannot be obtained at the phased array as it would require an incident zenith angle in air greater than 90 deg which is unphysical (cf. section 2.2.2). Therefore, these angles are limited to 35 deg. The circles around the stations estimated with these values are shown in Figure 25. Since the outer radius of the circle around stations 11 and 23 is large, the coordinates of the plot are limited such that the station and balloon positions can be distinguished (cf. Figure 26).

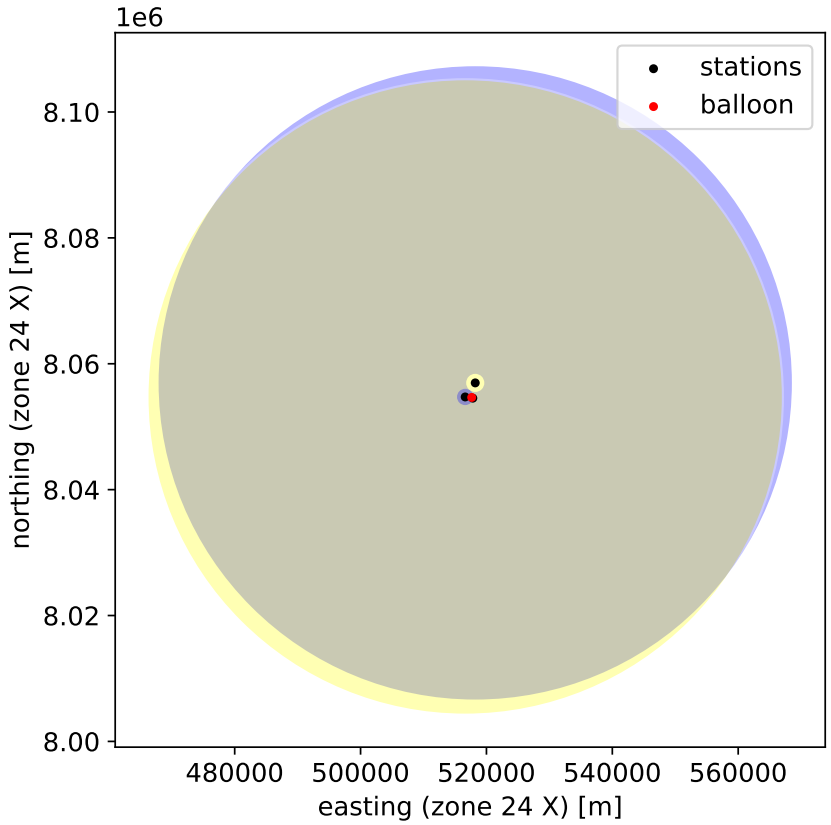


Fig. 25: Complete circular rings around stations denoting possible balloon locations. Additionally depicted is the balloon position from the GPS data. The width of the Gaussian peaks from the power distribution in dependence of the zenith angle is taken into account (radii of inner and outer circles). The circular ring in which the balloon is expected to be located is yellow for station 11. The circular ring corresponding to station 21 is magenta but hardly visible due to the large radii around station 11 and 23. And the circular ring where station 23 expects the balloon is blue. The area grayish brown encodes the superposition of the yellow and blue rings.

In Figure 26 (a), (b) and (c), the circular rings for the single stations are depicted separately including the GPS position of the radiosonde. The x- and y-coordinates are limited as mentioned before. Therefore, for station 11 and 23, only the inner circles and a colored area around them can be seen due to the outer radius exceeding the limits in the x- and y-direction. A superposition of the rings is shown in Figure 26 (d).

As can be seen in Figure 26 (d), the circular rings of stations 21 and 23 intersect at the position of the balloon but the radius of the inner circle of station 11 is larger than the distance between the balloon and station 11 in the xy-plane. This means that the balloon is not located on the cone above station 11, as it would have been expected. This is also shown in Figure 26 (a).

This observation can not be explained by the deviation between the marked and the actual balloon position at the time of data taking in station 11. A measurement in Station 11 is started three seconds after the time corresponding to the depicted balloon position. However, if an angle value for station 11 were accessible and displayed, the balloon position would still not be located in the corresponding circular ring. This is due to the absolute value of the zenith angle in station 11 decreasing in the time interval around 23:24 (cf. Appendix A). Therefore, the radius of the circular ring matching the real time of the value, i.e. three seconds earlier, in station 11 would be even larger than the one depicted in Figure 26. Accordingly, the angle of station 11 deviates from the expectation from the GPS position. This is consistent with the observation from section 4.

It is also to be considered that the radius of the inner circle of station 21, where the balloon is positioned, corresponds to an angle that is about one sigma away from the mean peak value of Figure 26 (a). It will be discussed in section 5.3.3 that these deviations are a limiting factor for the radiosonde tracking.

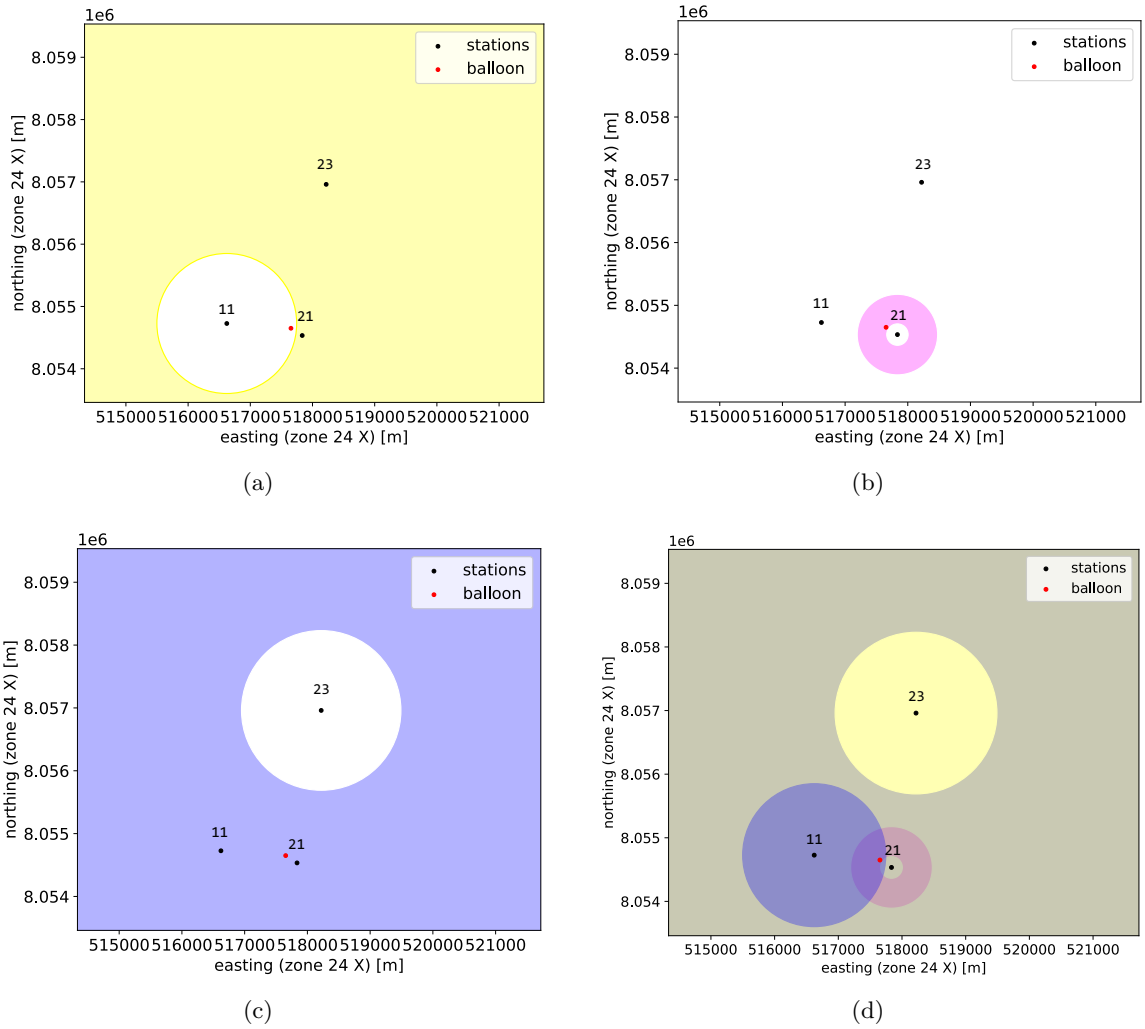


Fig. 26: Circular rings with limited x- and y-coordinates around station 11 (a), station 21 (b), station 23 (c) and a superposition of these rings (d). The GPS position of the balloon is additionally shown. The coordinates are limited to ensure a distinguishability of the single stations.

5.2 Ice Models

As described in section 5.1, the radius of the curved cone at the surface in ice ρ is calculated using a ray tracing algorithm. This algorithm depends on an ice model (cf. section 3.3). So far, a simplified ice model was used, as mentioned in section 3.3.

It is known that the single exponential function used for this model describes the refractive index at the surface and at the depth of the phased array well. However, the evolution of the refractive index between these depths is highly simplified and does therefore not represent the reality accurately [30]. A comparison of the refractive indices in dependence of the depth of the simple model (single exponential) and a more complex model 'greenland_poly5' (exponential polynomial) is shown in Figure 27.

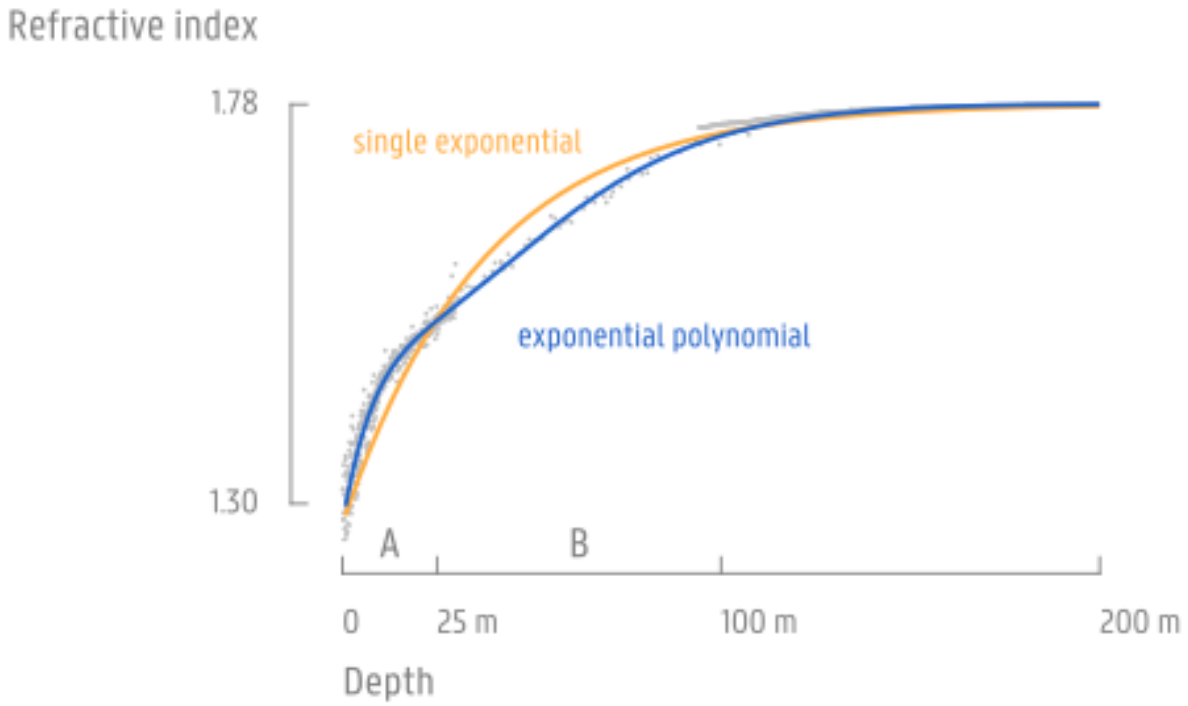


Fig. 27: Refractive indices of the Greenland ice sheet in dependence of the depth compared for two different ice models. Fig from Ref. [30].

To investigate the impact of the ice models on ρ, this property is depicted in dependence of the angle at the phased array for different ice models. Additionally to the ice models shown in Figure 27, an ice model based on a double exponential function called 'greenland_firn' is considered [14]. It is shown in Figure 28 that the difference between the ice models regarding the property ρ is in the order of one meter.

In section 5.3.2 it is discussed, that the uncertainty of ρ is in the order of tenth of meters. Therefore, the uncertainty due to the usage of the simple ice model is negligible. Thus, the simplest ice model 'greenland_simple' is used for the reconstruction of the balloon position (cf. section 5.3).

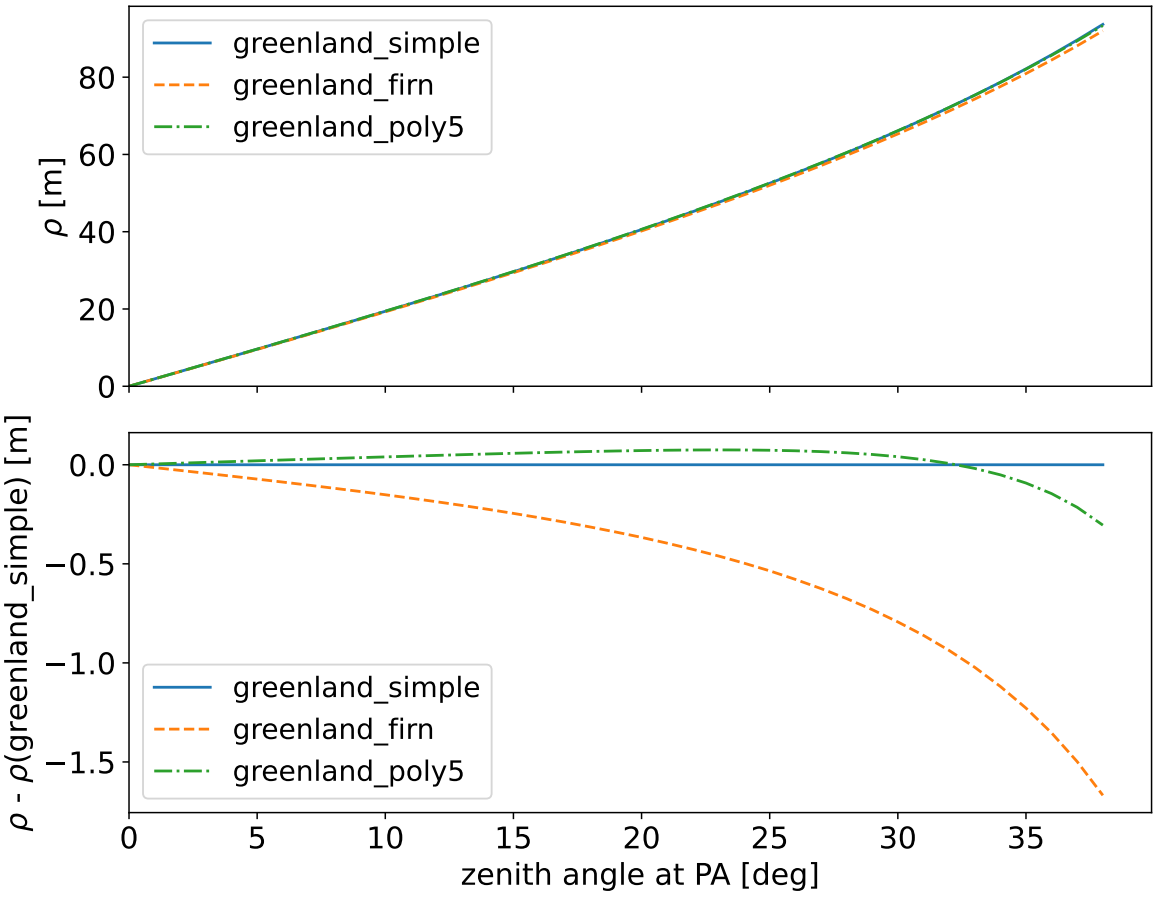
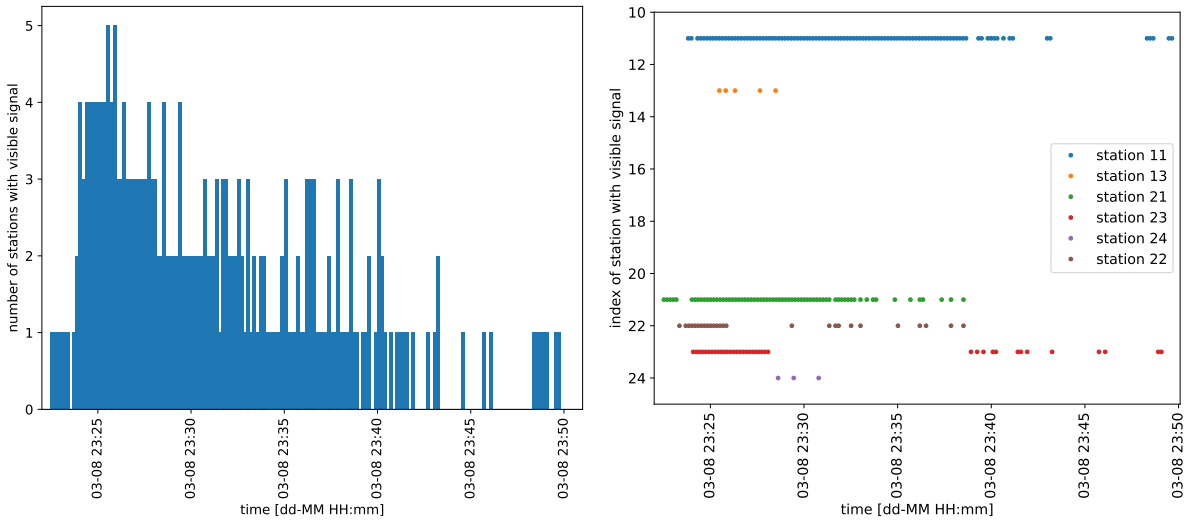


Fig. 28: Distance between station and surface interaction point ρ in dependence of the angle at the phased array (PA) for different ice models. Additionally depicted are the differences between the values of ρ and the value of the ρ from the simplest ice model used in this thesis.

5.3 Three Dimensional Reconstruction

In order to reconstruct the balloon tracks in three dimensions, at minimum three stations must measure a signal at the same time. As can be seen in Figure 13 and Figure 15, the stations do not show a continuous signal over the time of the balloon flight. It is therefore necessary to check that a sufficient amount of stations are yielding angles at the same time theoretically allowing for a reconstruction. An example of the number of stations showing a visible signal after phasing (cf. section 3.5.1) is shown in Figure 29. A signal at a station is considered visible, if the maximum power after the summation of the traces exceeds a threshold of 10 a.u.. A reconstruction is not possible at every time, but there are still some time intervals where three or more stations have a visible signal.



(a) (b)

Fig. 29: Number of stations showing a visible signal (after phasing) over time as histogram (a) and per station index (b).

5.3.1 Fitting Method

An option to reconstruct the balloon position is to use a χ^2 minimization. This minimization is based on a χ^2 function. In general, such a function is of the following form [31]:

$$\chi^2 = \sum_{i,j}(x_i - y_i(a))V_{ij}(x_j - y_j(a)) = \sum_i \frac{(x_i - y_i(a))^2}{\sigma_i^2} \tag{10}$$

In this function x is a vector containing the observations and $y(a)$ is a vectorial expression of the fitted values in dependence on the fit parameters a. In the case of reconstruction of the radiosonde, the fit parameters are the coordinates x_b, y_b and z_b that should be reconstructed. The property V is the inverse error matrix corresponding to the observations in the vector x and σ^2 is the corresponding inverse diagonal element of V, which is expected to follow a Gaussian distribution. Note that the last equality only applies when statistical independence of the observations is given, i.e. when the matrix V is diagonal [31].

An important part of a fit is also to determine its goodness and the uncertainties of the different fit parameters. The `iminuit` python interface (cf. [32]) is used for the minimization in this thesis. It is based on the C++ tool `MINUIT` and provides multiple techniques to estimate the uncertainties of the fit parameters. It is to be kept in mind that in all these techniques, a proper normalization of the χ^2 function by σ_i is crucial. An incorrect normalization may lead to misinterpretation of the provided uncertainties. For example, if the uncertainty σ_i is overestimated by a certain factor, the uncertainty returned by Minuit is overestimated by the same factor. In `MINUIT` an internal normalization factor called UP is defined which is by default set to one. A change in the parameter value resulting in a modification of the function value equal to the factor UP is set as the uncertainty of the variable. For UP = 1 this returns one-standard-deviation uncertainties. When two-standard-deviation uncertainties are needed UP should be set to four due to the quadratic behavior of the function. Subsequently, only one-standard-deviation uncertainties are considered. It is common in statistics to define an uncertainty based on the second derivative of the χ^2 function. MINUIT realizes this by the HESSE technique. In case of a parabolic behavior at the minimum, the `HESSE` algorithm leads to the same values as the uncertainty coming from searching the parameter value at which the change in the χ^2 function

equals UP [31].

A sketch of the full trigonometric case of the balloon coordinates (x_b, y_b, z_b) in relation to the phased array coordinates (x_{pa}, y_{pa}, z_{pa}) and the position of the surface interaction point (x_s, y_s, z_s) is shown in Figure 30. The lower triangle including the phased array position and the position of the surface interaction point is only an approximation. In reality, the connecting line between the two points is curved and the distance ρ in the xy-plane is calculated through ray tracing.

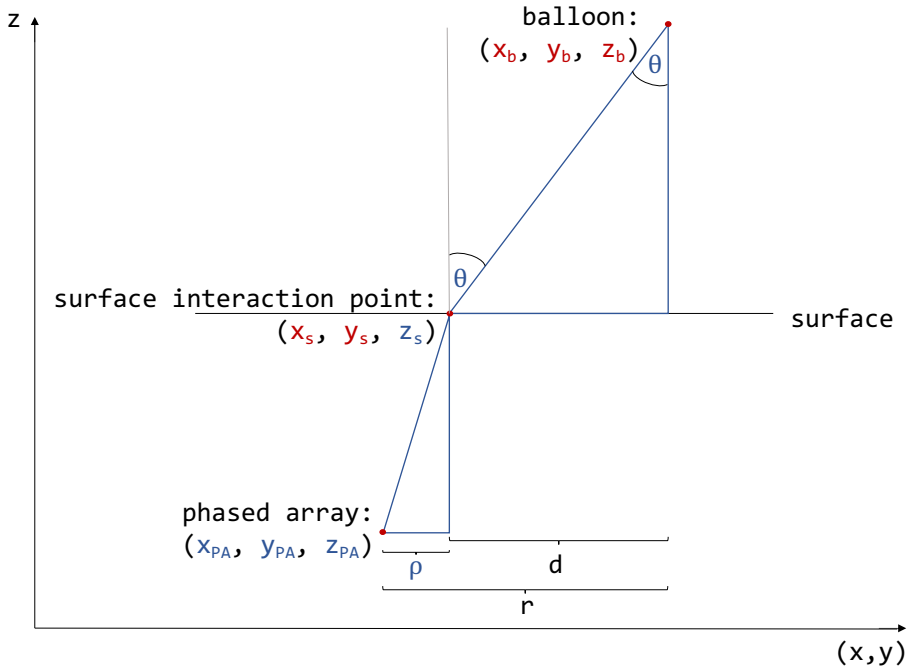


Fig. 30: Sketch of trigonometric case of balloon coordinates in relation to phased array coordinates and position of surface interaction point in z over (x,y) coordinates. Known variables are colored blue, whereas unknown variables are shown in red. The properties in black are used to shorten the obtained function.

In practice, the χ^2 function used in the minimization of the balloon position, based on these considerations, has a slightly more involved expression with respect to Equation 10. It is given by:

$$\chi^2 = \sum_{i=1}^{n} \left(\frac{\tan(\theta_i) \cdot (z_b - z_s) + \rho_i - \sqrt{(x_b - x_{pa,i})^2 + (y_b - y_{pa,i})^2}}{\sigma_i} \right)^2 = \sum_{i=1}^{n} \left(\frac{d_i + \rho_i - r_i}{\sigma_i} \right)^2 \tag{11}$$

where n is the total number of used stations, d is the distance between the surface interaction point and the radiosonde position (xy-plane). The distance between station and surface interaction point (xy-plane) is given by the parameter ρ provided by the ray tracing process and r is the distance between the station and the balloon position also given in the xy-plane. The numerator $d + \rho - r$ of Equation 11 will be referred to as F.

Three stations are needed to obtain a determined system of equations, as this leads to three terms in the χ^2 function for three variables that are to be determined. An overdetermined system is generated when using four or more stations. A χ^2 minimization is usually based on

an overdetermined system as this improves the error estimation and enables the determination of the goodness of the fit. A measure for the goodness of fit is given by the so-called reduced χ^2 (χ^2_{red}) which is defined as [33]:

$$\chi^2_{\mathrm{red}} = \frac{\chi^2}{\mathrm{N_{DoF}}} \tag{12}$$

The degrees of freedom $\mathrm{N_{DoF}}$ of the system are given by the difference between the number of fit parameters and the number of equations. For data that does not match the model, χ^2_{red} is larger than one. Is the value of χ^2_{red} less than one, the model is regarded as overfitted [33]. Therefore, for a good fit, the χ^2_{red} of a single data point is expected to be approximately one and the average is expected to yield $\langle\chi^2_{\mathrm{red}}\rangle = 1$ [33].

The calculation of the σ_{i} is described in detail in section 5.3.2.

5.3.2 Error Calculation

A χ^2 function as described in section 5.3.1 includes a weighting factor σ_{i} for its single terms.

Applying Gaussian error propagation directly to the numerator F of Equation 11 would result in a lengthy term even for a single station. Therefore, at first the process of calculating the uncertainty is divided into smaller steps as subsequently demonstrated.

Mathematically the properties d and r of F can be calculated by:

$$\begin{aligned} \mathrm{d} &= \tan(\theta) \cdot (\mathrm{z_b} - \mathrm{z_s}) \\ \mathrm{r} &= \sqrt{(\mathrm{x_b} - \mathrm{x_{pa}})^2 + (\mathrm{y_b} - \mathrm{y_{pa}})^2} \end{aligned} \tag{13}$$

It is assumed that the uncertainties of the x- and y-coordinates of the phased array ($\mathrm{x_{pa}}$,$\mathrm{x_{pa}}$) are negligible. In case of a χ^2 function, the fit parameters are assumed to be exactly known, i.e. they have no uncertainties. Therefore, the term r has no uncertainties in this calculation.

Lower and upper bounds for the uncertainties of ρ and θ can be estimated using the ray tracing technique and taking the zenith angle α at the phased array as starting point for the calculation. Let $\alpha - \sigma_\alpha$ be the lower and $\alpha + \sigma_\alpha$ be the upper bounds of the uncertainties of α. Where σ_α is the width of the Gaussian peak (cf. section 5.1). Ray tracing is applied to the angles $\alpha - \sigma_\alpha$, α and $\alpha + \sigma_\alpha$ through which the values for ρ_{low}, θ_{low}, ρ, θ, ρ_{up} and θ_{up} are obtained. The upper and lower bound for the uncertainties of ρ are therefore given by $\Delta\rho_{\mathrm{low}} = \rho - \rho_{\mathrm{low}}$, respectively $\Delta\rho_{\mathrm{up}} = \rho_{\mathrm{up}} - \rho$. The uncertainties for the angle θ at the surface in air can be expressed analogously. An edge case occurs if $\alpha - \sigma_\alpha < 0\,\mathrm{deg}$, respectively $\alpha + \sigma_\alpha > 35\,\mathrm{deg}$ as for these values the corresponding surface interaction point and the surface angle can not be estimated (cf. section 2.2.2). In this case, the ray tracing is performed with angle $\alpha - \sigma_\alpha$ assigned to zero deg, respectively $\alpha + \sigma_\alpha$ assigned to 35 deg.

The uncertainty of the property d can then be calculated using Gaussian error propagation:

$$\Delta\mathrm{d} = \sqrt{(\tan(\theta) \cdot \Delta\mathrm{z_s})^2 + \left(\frac{\mathrm{z_b} - \mathrm{z_s}}{\cos(\theta)^2} \cdot \Delta\theta\right)^2}$$

During the fit process the values of the σ_i are updated in each iteration. To track the evolution of the uncertainties in the individual substeps, an example calculation is carried out with the

GPS-values of the balloon position in the following. In section 3.2 it was explained that the elevation of the surface observed in the radiosonde data deviates in the order of tenth of meters from the value from $z_s = 3251\,\mathrm{m}$. To include the deviation of the alignment of the two coordinate systems in the calculation, the uncertainty of z_s is assumed to be 10 m.

As can be seen from Figure 31, considering station 23, the uncertainties in the order of a few degrees of the zenith angle at the phased array result in large uncertainties in the order of thousands of meters of the distance d. As F is linear in the terms of r, d and ρ, a large uncertainty in one of these terms results in a large overall uncertainty of the function.

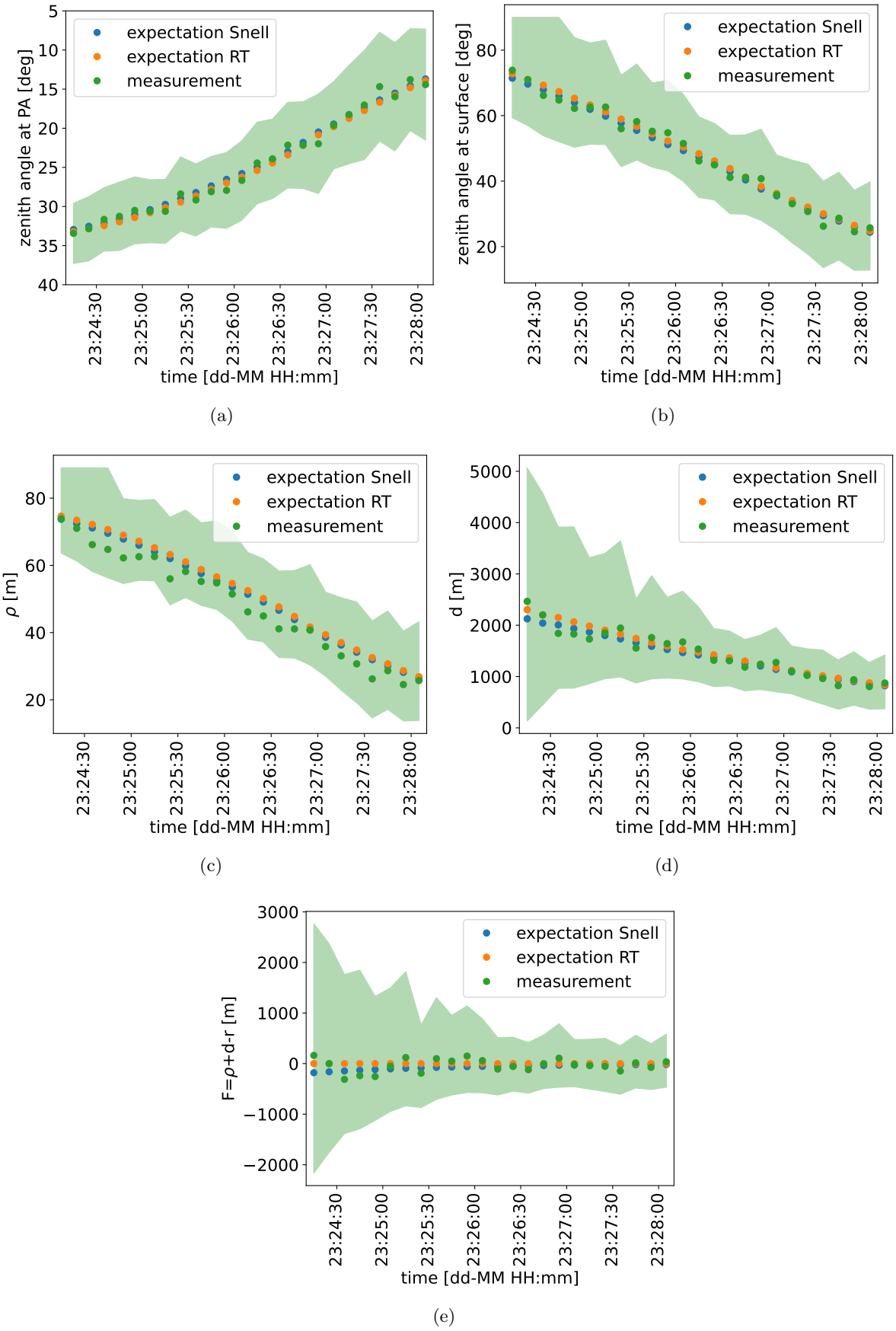
The calculation of F, using the expectation from the ray tracing procedure for station 23, leads to values that are approximately zero. This is expected, as r and ρ+d describe the same distance (cf. Figure 30).

In section 3.3 it was discussed that the difference between the expected angular values at the phased array from the trigonometric calculation and the ray tracing estimation is small. Therefore, it is expected that the F values determined by using the trigonometric calculation are also near zero. It is shown in Figure 31 (e), that the values of F based on the trigonometric expectation deviate from the values of F from the ray tracing expectation and thus deviate from zero for the first data points. This deviation is also observed in the property d (cf. Figure 31 (d)), which is therefore causing the obtained deviation in F.

In general, the deviations between the ray tracing expectation and the measured angles at the phased array are larger than the deviations between the different expectations. Therefore, a deviation in the values of F for the measurement from zero is observed (Figure 31 (e)). This is a limiting factor for the reconstruction of the weather balloon, as the χ^2 function based on F is used for the fitting process.

The deviation is explained by the large uncertainties from the Gaussian error propagation. However, it is possible that the uncertainties are overestimated, as they are in the order of thousands of meters, which is large compared to the observed fluctuations in the values of F. This may be a result of overestimated uncertainties of the zenith angle at the phased array. As explained, the one-standard-deviation-uncertainties of the Gaussian fits (cf. section 5.1) are used as starting point for this error propagation. The signals measured by the different phased array antennas are not longer fully coherent for an angle deviating by σ_α from the angle α. Thus, the uncertainty of α used in this calculation may be overestimated.

The uncertainties σ_i needed to weight the χ^2 function from Equation 11 can not be provided to the function as lower and upper bounds but have to be a symmetric value for the whole term of one station. Thus, the upper and lower limits of the errors are averaged to generate only one value for σ_i. This is only an approximation.



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

Fig. 31: Single steps of Gaussian error propagation for function F taking the asymmetry of the situation into account. Depicted are the time evolution for the angle α at the phased array (a), the angle θ at the surface (b), ρ (c), d (d) and F (e). Uncertainties for the values coming from the measurement are illustrated by the green bands. Station 23 is considered. RT is short for ray tracing and PA is the abbreviation for phased array.

5.3.3 Results

In this section, the fit results for different station combinations are discussed. This includes fits with different combinations of three stations, as well as a fit based on four stations. Furthermore, the uncertainties of the fit parameters provided by the fit and calculated as described in section 5.3.2 are compared. Finally, a fit using the expected angular values at the phased array is investigated.

The fit results for the χ^2 function given in Equation 11 for different combinations of three channels each are depicted in Figure 32. The parameters ρ_i and θ_i are calculated through the ray tracing procedure applied to the mean values of the Gaussian fits of the power peaks, as described in section 5.1. The weighting factors σ_i are estimated separately, as explained in section 5.3.2, using the known GPS values of the balloon position. Furthermore, times for the fit were selected where the same stations, in this case stations 11, 21, 22 and 23, show a continuous signal (cf. Figure 29). Data points measured within an interval of ten seconds are considered as taken at the same time (cf. section 5.1). The x- and y-coordinates of station 21 and a balloon altitude of $z_b = 0$ are used as initial parameters provided to the fit.

Figure 32 shows that the slope of the time evolution of the northing (y-coordinate) and the altitude (z-coordinate) of the fitted values approximately match the expectation (GPS track) for all channel combinations. For the easting (x-coordinate), the time evolution of the fit values does not match the GPS track evolution as well as the other coordinates.

Furthermore, it is apparent that the choice of stations used for the fit changes the absolute values of the fitted parameters. For the easting, the time evolution of the data points is similar for all fits considering station 11. Only the shape for the combination of stations 21, 22 and 23 deviates from the others. This indicates that station 11 has a larger influence on the evolution of the data points in the x-direction than the other stations. Subsequently, it is discussed that this might be a general issue of the station arrangement.

In order to verify that the fit is producing reasonable values for the fit parameters, especially for the x-coordinate, again the two dimensional case is considered. As described in section 5.1 the circular rings for station 11, 21 and 23 do not intersect at the time of the first point in the fit (Figure 32) as the inner radius for station 11 is too large (cf. Figure 26). This already indicated potential issues during the fit.

For further analysis of the first data point of the fit (channel combination 11, 21 and 23) the elevation of the balloon is fixed to the GPS value at that time. In Figure 33 (a), (b) and (c), circles are depicted, representing the radii of the top circular bases of the respective cones produced by possible zenith angles at the phased array (cf. section 5.1). The angles are chosen in one degree steps. The largest circle corresponds to the largest angle considered for calculating the coherent sum (cf. section 3.5.1) that can be processed by the ray tracing technique with subsequent refraction into air.

The circles in Figure 33 (a), (b) and (c) are colored according to the corresponding power value. The color of the larger circle is furthermore assigned to the space between this circle and the next smaller one. Additionally, the radii $\rho_i + d_i$ generated by the mean value of the Gaussian fit of the respective station is also visualized. The GPS position and the reconstructed position of the balloon are shown. The values of the reconstructed position depicted are the result of a fit of the first data point with fixed z-coordinate. This is necessary as the z-coordinate of the fit shown in Figure 32 deviates from the expectation and would cause inconsistencies in the interpretation of the plot. The distances of the fit values from the respective station positions correspond to the quantity r_i from Equation 11.

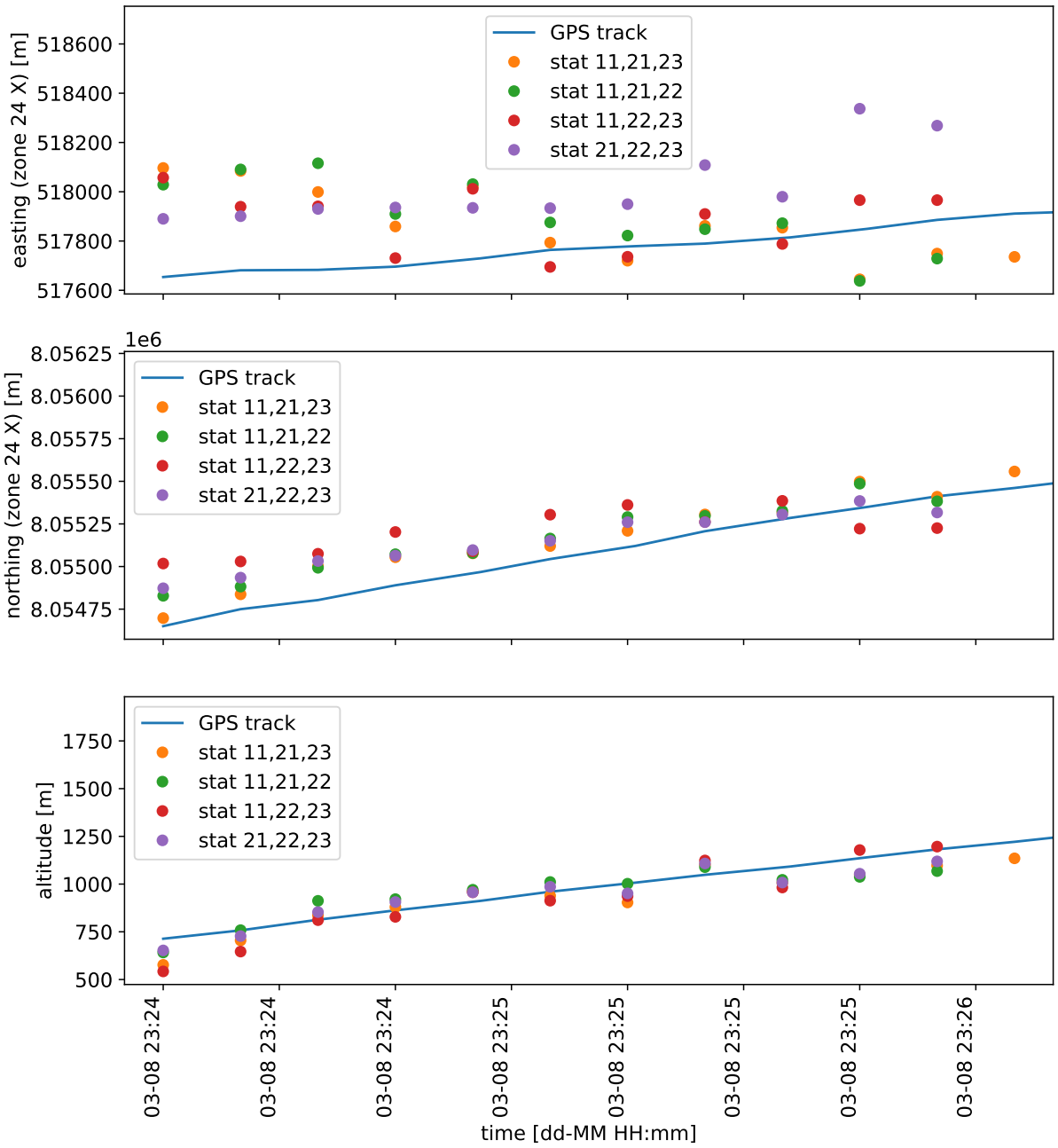


Fig. 32: Fit results for different channel combinations of three stations each showing reasonable signals with weighting of the χ^2 function in comparison to the GPS position of the balloon.

As described in section 5.3.1, the goal of the fit is to minimize the distance between $\rho_i + \mathrm{d}_i$ and r_i.

In Figure 33 (a) the situation for station 11 is shown. It can be seen that the GPS position of the balloon corresponds to an angular value to which a lower power value is assigned. This is equivalent to the observation in Figure 26. There it was shown that the radius of a circle corresponding to the lower one-standard-deviation angle is larger than the distance between the station and the GPS position of the radiosonde.

From Figure 33 (a) it is furthermore apparent, that the fit yields a reconstruction of the position that deviates less from the circle with radius $\rho_{11} + \mathrm{d}_{11}$. This is expected, as the fit algorithm minimizes the difference between the property $\mathrm{r}_{\mathrm{fit}}$ (including the coordinates $\mathrm{x_b}$ and $\mathrm{y_b}$) and the fixed value of $\rho_{11} + \mathrm{d}_{11}$.

For stations 21 and 23 (Figure 33 (b) and (c)) it can also be seen that the fit minimizes the distance between $\mathrm{r}_{\mathrm{fit}}$ and $\rho_{21} + \mathrm{d}_{21}$, respectively $\rho_{23} + \mathrm{d}_{23}$. This behavior is again expected.

In Figure 33 (d), the resulting χ^2 values for the combination of stations 11, 21 and 23 are represented. For better visibility, the values of the χ^2 function are scaled using the natural logarithm. It is shown that the reconstruction in x-direction is particularly affected by the deviation of the angle value in station 11. It can be seen that the circles with radii $\rho_{21} + \mathrm{d}_{21}$ and $\rho_{23} + \mathrm{d}_{23}$ intersect twice near station 21. The GPS position of the balloon is located near the intersection with smaller values in the easting. However, the fit reconstructs the position of the balloon near the other intersection, as the value of the χ^2 function is smaller at this point. Due to the arrangement of the stations, there is no veto in x-direction against the too large radius of station 11, resulting in χ^2 values smallest near the reconstructed balloon position.

The arrangement of the stations has a great influence on the result of the fit when only a few stations are used. Assume that all stations are arranged in a row. In an ideal case, i.e. when all these stations show an angle matching the expected angle between station and balloon, the top circular bases of the cones above the stations (cf. section 5.1) intersect twice. These intersections are located symmetrically around the connecting line between the stations. If the stations do not show an ideal signal, there are multiple intersections of two top circular bases each. These intersections are still distributed symmetrically around the connecting line between the stations. Therefore, a fit using only this kind of arrangement in two dimensions is most likely to result in an under-determined system as no intersection area is favored. When looking at the correlation matrix of the three dimensional fit with stations 21, 22 and 23 for the penultimate data point (cf. Table 2), it can be seen that the x- and y-coordinates are maximally anti-correlated, leading to the fact that this fit is only approximate.

Tab. 2: Covariance correlation matrix for fit with stations 21, 22 and 23

	$\mathrm{x_b}$	$\mathrm{y_b}$	$\mathrm{z_b}$
$\mathrm{x_b}$	1	−1	0.8
$\mathrm{y_b}$	−1	1	−0.8
$\mathrm{z_b}$	0.8	−0.8	1

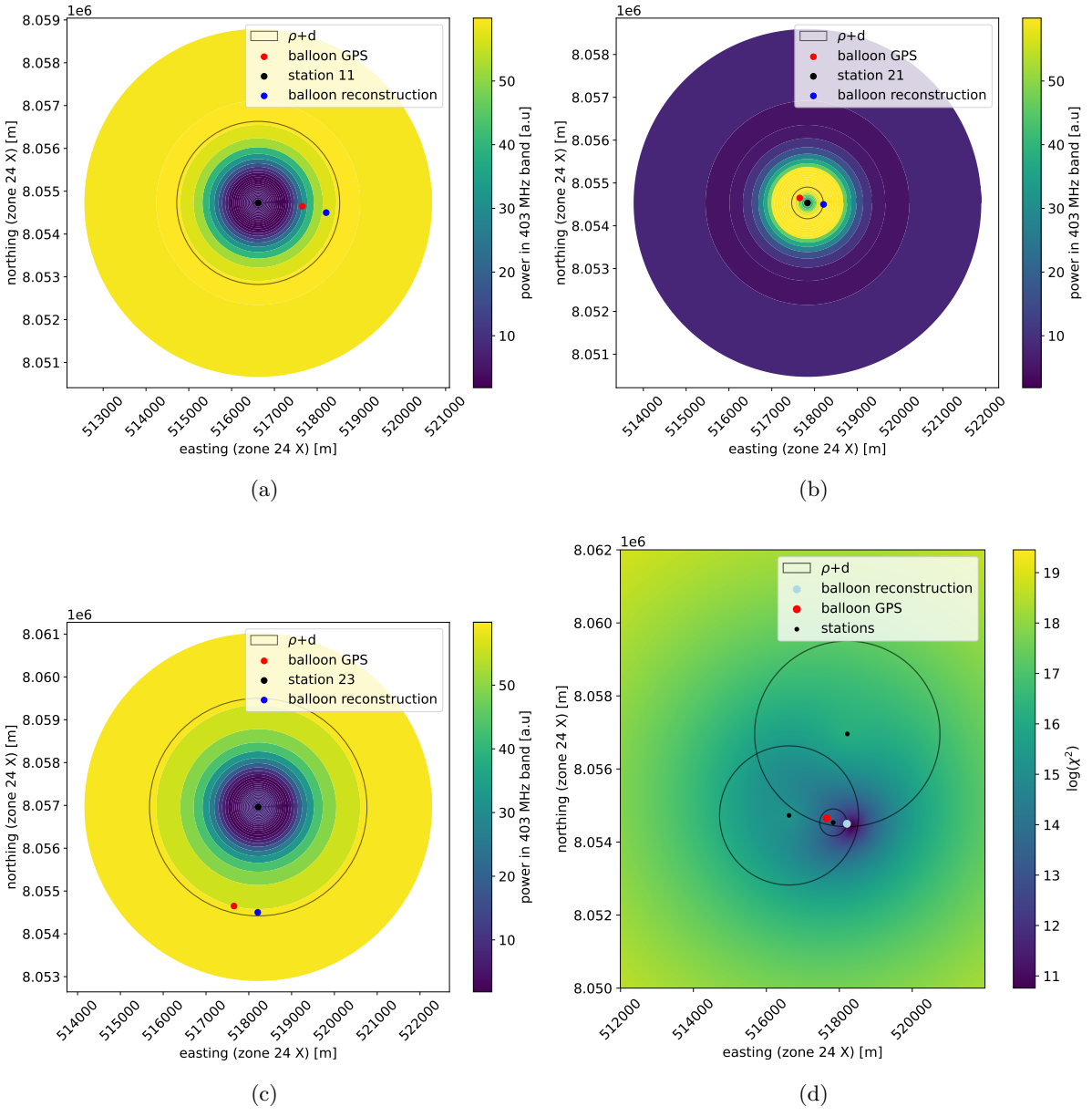


Fig. 33: Circles above station 11 (a), 21 (b), 23 (c). The color in these three plots encodes the power in the 403 MHz band. In (d) the color encodes the natural logarithmic values of the χ^2 for these stations in dependence of x- and y-coordinates. Additionally, the balloon positions from the GPS data and the fit for a fixed z-coordinate are depicted.

Now assume that all stations are arranged in a row except for one. The intersection areas of the stations in a row are again distributed symmetrically around their connecting line. The radius of the other station (the one not included in the row) is responsible on which side of the connecting line the balloon position is reconstructed. Therefore, a too large, respectively too small, radius at this station leads to inaccuracies in the reconstruction.

This issue is also present in the full three-dimensional fitting case.

In section 5.3.4, a different fit approach is discussed which, under certain conditions, provides a more accurate reconstruction of the x-coordinate. This fit takes an expression for the x- and y-coordinates of the surface interaction points into account. This allows the fit to choose the correct point of intersection.

It is expected that using four or more stations will lead to better fit results, as stations with higher uncertainties are given less weight in the overall fit. In addition, determining the goodness of the fit is only possible through taking more than three stations into account as mentioned in section 5.3.1. Therefore, in Figure 34 the fit results for four stations are compared to the results of a fit using only station 11, 21 and 23. Again, the x- and y-coordinates of station 21 and a balloon altitude of $z_b = 0$ are used as initial parameters provided to the fit.

From Figure 34 can be seen that the fit including four stations looks similar to the fit using only stations 11, 21 and 23 and does not result in a significantly better outcome. Also for 4 stations the evolution of the fit values in x-direction does not correspond to the development of the GPS track. Therefore, the expectation of a better reconstruction is not met.

As described in section 5.3.1 a perfect fit with one free parameter would result in $\langle \chi^2_{\text{red}} \rangle = 1$. For a single data point, this value will not be reached. Instead, for data points matching the expectation of a model, the reduced χ^2 is expected to reach values near one. The χ^2_{red} values for the four station fit case are listed in Table 3. It can be seen that these values are much smaller than one. This case usually occurs when the values of σ_i (weighting factors) are overestimated or the model is too flexible and therefore small fluctuations have an influence on the fit result [33].

Furthermore, in the case of the four station fit, the σ_i of the χ^2 function are calculated as described in section 5.3.2. Based on this calculation, the underlying uncertainties for this property are not symmetric and therefore not Gaussian distributed (cf. section 5.3.2). This may lead to the observed behavior of the χ^2_{red}. Usually, in the case of not Gaussian distributed uncertainties, a χ^2 minimization should not be used [33].

Tab. 3: Values of χ^2_{red} for the different data points used for the four station fit. Index denotes the position of the data point in the reconstruction fit.

index	0	1	2	3	4	5	6	7	8	9	10
χ^2_{red}	0.10	0.03	0.007	0.007	0.00019	0.019	0.04	0.0010	0.0012	0.0006	0.006

A further analysis of the deviations of the fit values from the GPS values is shown in Figure 35. For the x-, y- and z-direction the difference of the GPS and the fit values is depicted in a histogram. In this plot it is shown even more clearly, that the reconstruction of the x-coordinate has the largest deviations from the expectation, reaching values near 400 m, whereas the reconstruction of the balloon altitude yields the smallest deviations. Furthermore, the differences of the z-direction are distributed symmetrically around zero. For the y-coordinate, the fit produces only values that match the expectation or are larger than the expectation.

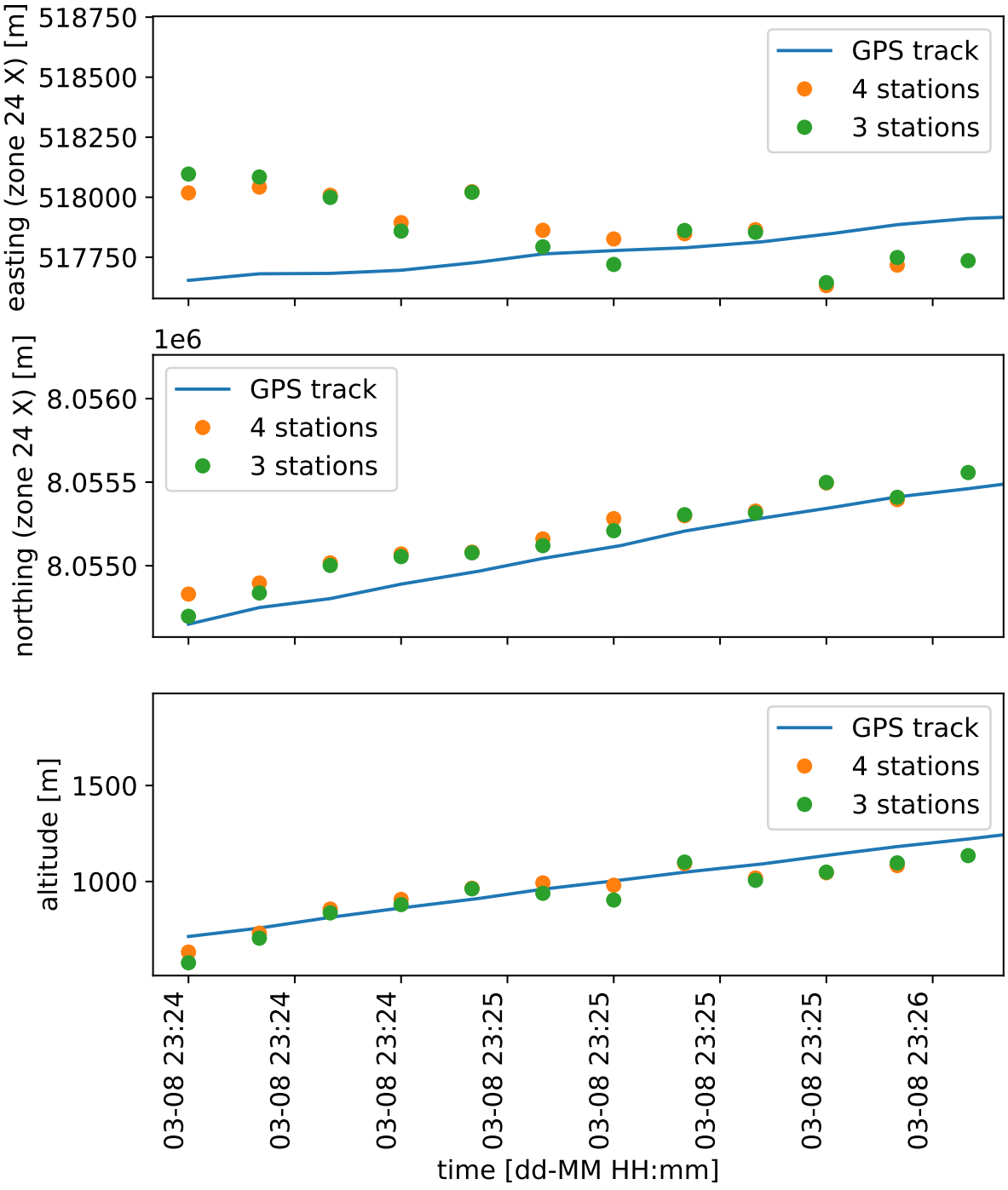


Fig. 34: Fit results for three (11,21,23) and four stations (11,21,22,23) showing reasonable signals with weighting of the χ^2 function in comparison to the GPS position of the balloon .

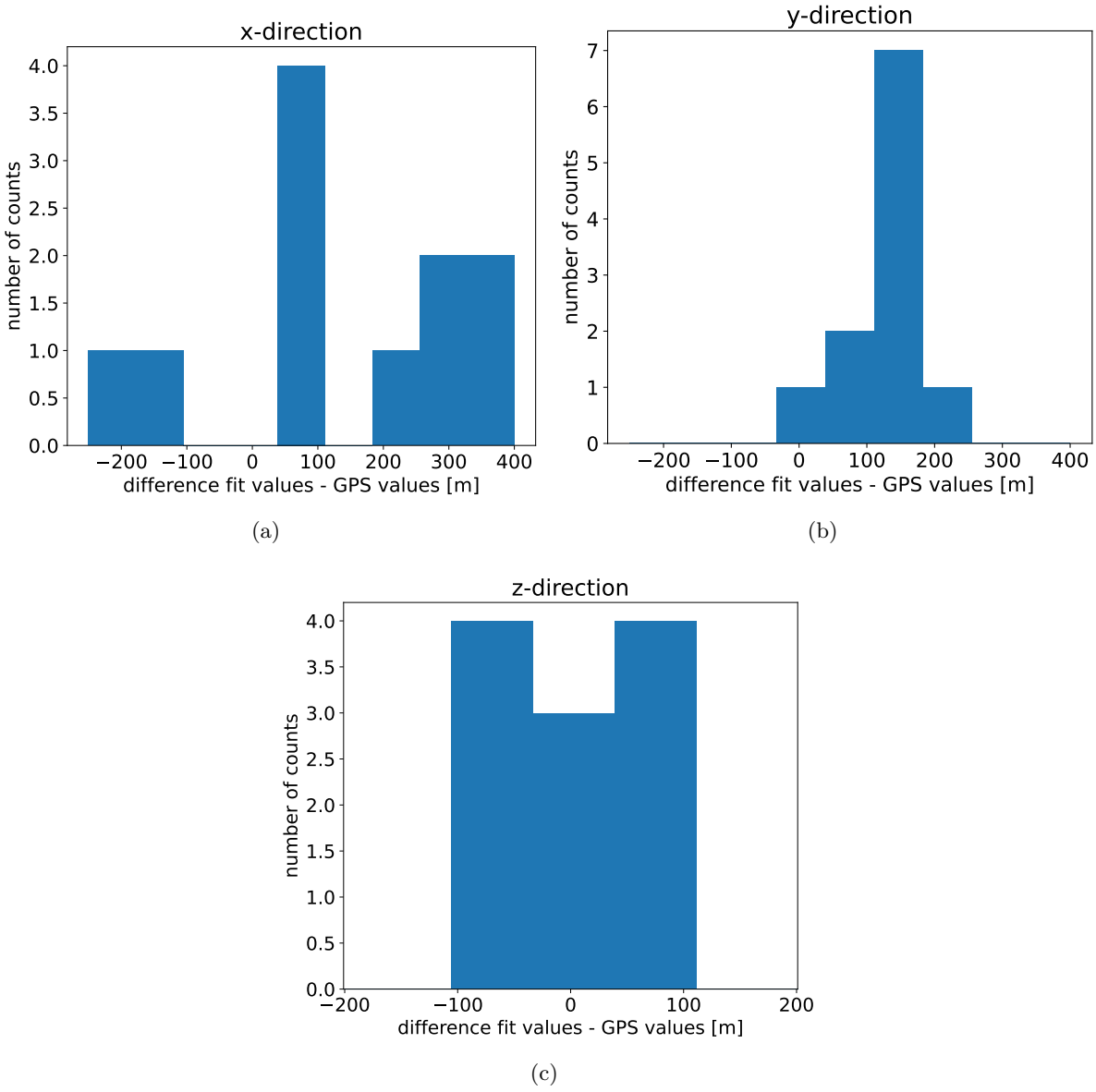


Fig. 35: Histograms of differences between fit values and GPS data for x-direction (a), y-direction (b) and z-direction (c).

The issue in the x-coordinate reconstruction is most likely caused again by the values of station 11 being not explicitly vetoed for as illustrated by Figure 33.

To get a more accurate fit result, it would be useful to include a station on the other side of station 21 in the x-direction (opposite station 11). This is not possible at this time because the stations in question are not yet operational respectively have yet to be built. In general, it is assumed that fitting works best when many stations with as many different x- and y- positions as possible are used.

Additionally, it is expected that a fit using data points actually measured at the same time would lead to more accurate results. Therefore, it is reasonable to perform a fit taking all triggered events into account. This would have been beyond the scope of this thesis.

In section 5.3.1 it is described that `iminuit` also provides an estimation of the uncertainties of the fit parameters. Furthermore, it is mentioned that a correct normalization σ_i, i.e. weighting

of the χ^2 function, is necessary to provide reasonable uncertainties. To examine what effect the weighting has on the uncertainties, different estimations are compared in Figure 36.

On the one hand, the calculation of the σ_i is directly included in the fit as the fit values are part of the expression of the Gaussian error propagation (cf. section 5.3.2). The results including the uncertainties provided by `iminuit` are shown in Figure 36. On the other hand, the results for a fit are shown where the σ_i are calculated separately using the GPS values. Additionally, a fit with $\sigma_i = 1$ together with the corresponding uncertainty is depicted. It can be seen that only the weighted fits are providing reasonable uncertainties. The uncertainties of the fit without weighting are not sufficient to explain the deviation of the fit values from the GPS values. Furthermore, it can be seen that the uncertainties provided by the weighted fits are similar. Therefore, both methods can be used equally. The uncertainties provided by both weighted fits are able to explain the observed deviation from the expected positions. However, the uncertainties are large compared to the evolution of the GPS expectation. This indicates, together with the observations of the χ^2_{red}, that the uncertainties calculated as described in section 5.3.2 are overestimated.

Additionally, it can be seen that the different weighting factors only have a small influence on the absolute fit values. The observed deviation most likely comes from the different absolute values that the χ^2 function takes during the fit. Depending on the absolute values, the termination condition of the algorithm is reached sooner or later. Therefore, a not weighted fit is sufficient if the uncertainties and the goodness of the fit are not of interest.

In order to verify that the issues encountered by the fit indeed are based on imprecise angle values measured by the phased array, the fit is performed with the expected angular values at the phased array. In Figure 37 the results are displayed for a fit using the expectation calculated trigonometrically via Snell's law and the ray tracing process (cf. section 3.3). Again, the x- and y-coordinates of station 21 and a z-coordinate of zero are considered as initial values for the fit. For simplicity, the weighting factor is set to one. For looking at nominal values and not considering the uncertainties, this simplification is justified as mentioned before.

Taking Figure 37, it can be seen that the values of the fit using the ray tracing expectation match the GPS track. This is expected as the function used to calculate the ray tracing expectation is similar to the function used by the fit (cf. section 3.3). Fitting the expectation calculated trigonometrically using Snell's law (assuming a depth of the phased array of -100 m) results in matching values for the easting and northing but a small deviation in the altitude reconstruction. Therefore, it can be seen that the fit is sensitive to small deviations in the angle at the phased array.

Therefore, in order to obtain better fitting results, it is necessary to optimize the angle reconstruction in the phased array. This includes a more granular sampling of the possible delays in the phasing process and an optimization of the determination of the resulting angle. So far, a Gaussian fit was used. Integrating a gain calibration into the weighting of the fit and checking for other sources of error (e.g., cable delays) can also contribute to a better fit as well as choosing an other station arrangement when possible. Additionally, considering more balloon flights allows to verify reproducibility of the described observations.

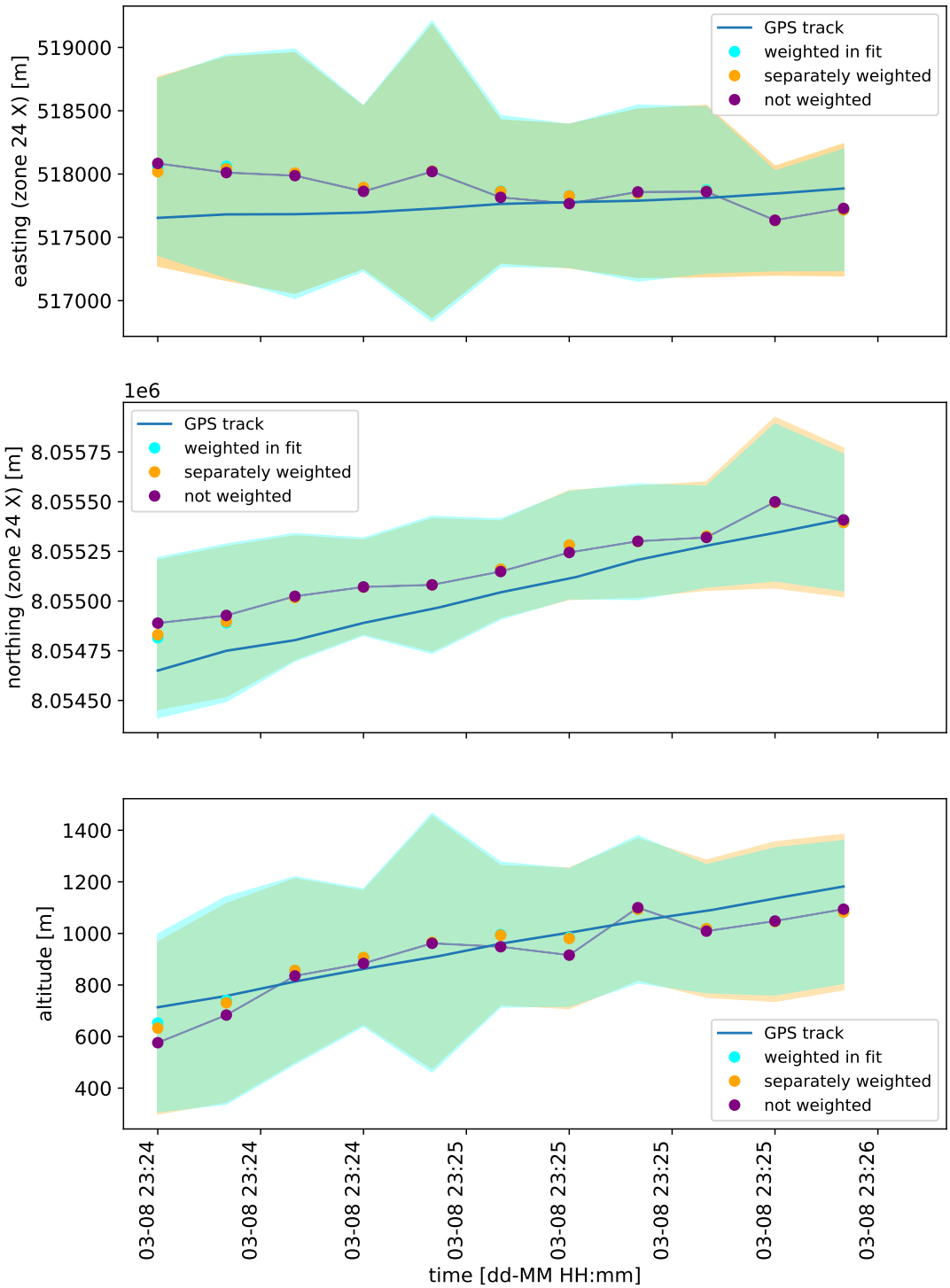


Fig. 36: Fit results for four stations (11,21,22,23) with and without weighting in comparison to the GPS position of the balloon. Errors from `iminuit` output. Different weighting techniques are used (separated calculation and calculation integrated in fit).

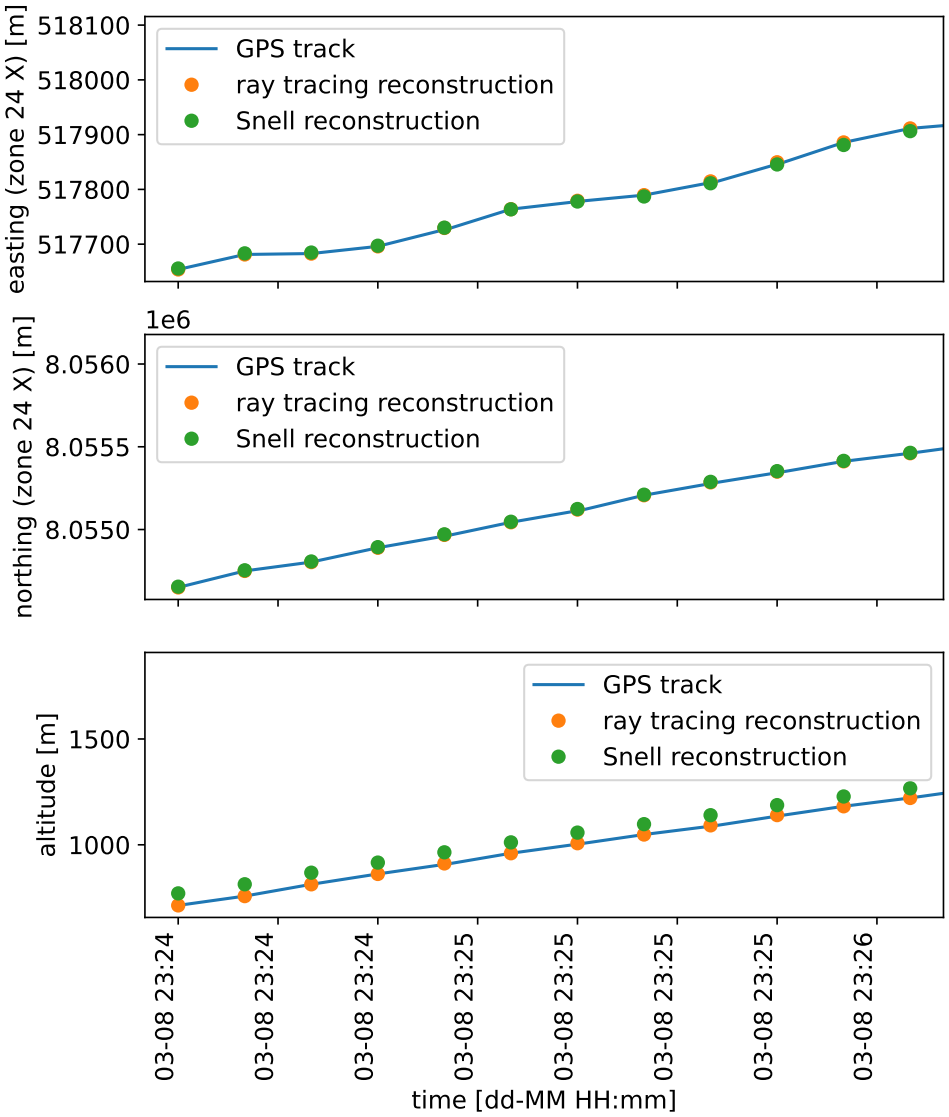


Fig. 37: Fit results for four stations (11,21,22,23) with expected angle values from ray tracing and trigonometric calculation using Snell's law.

5.3.4 Fit with Auxiliary Variables

In the analysis described in section 5.3.3, the reconstruction of the x-coordinate is deviating more from the expectation as the other coordinates. This was attributed to inaccuracies in the angular values of station 11 in combination with the station arrangement. In this section, a fit based on a more complex χ^2 function is described. This function takes the position of the surface interaction point into account. Therefore, it is expected that this fit improves the reconstruction of the x-component.

A sketch of the trigonometric case of the balloon coordinates in relation to the phased array coordinates and the position of the surface interaction point in the xy-plane is shown in Figure 38. The underlying structure of the full three dimensional case is the same as depicted in Figure 30.

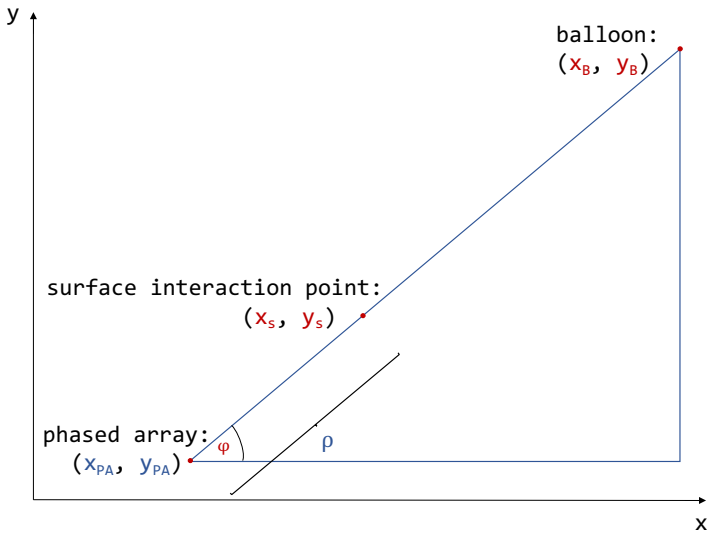


Fig. 38: Sketch of trigonometric case of balloon coordinates in relation to phased array coordinates and position of surface interaction point as projection on the xy-plane. Known properties are marked blue, whereas red colored variables are unknown for the fitting process.

Based on these considerations, the relation between the balloon position (x_b, y_b, z_b), the coordinates of the surface interaction point (x_s, y_s, z_s) and the surface angle of the radio waves θ for one station can be described by using the following equation:

$$\tan(\theta) = \frac{\sqrt{(x_b - x_s)^2 + (y_b - y_s)^2}}{z_b - z_s} \tag{14}$$

This basic equation would be sufficient to reconstruct the balloon position when using three stations for the reconstruction, if all coordinates of the surface interaction point were known. However, for the surface interaction point, only the z-coordinate z_s and the distance ρ to the phased array are known or can respectively be calculated starting from the observed angle at the phased array (cf. section 5.1). Therefore, a second equation is required that considers the angle ϕ between the x- and y-components of the line connecting the station and the surface interaction point with the balloon (cf. Figure 38). The following equations apply:

$$\begin{aligned} x_s &= \rho \cdot \cos(\phi) + x_{PA} \\ y_s &= \rho \cdot \sin(\phi) + y_{PA} \\ \tan\phi &= \frac{y_b - y_{PA}}{x_b - x_{PA}} \end{aligned} \tag{15}$$

Combining Equation 14 and the system of Equations 15 and inserting weighting factors $w_{1,i}$ and $w_{2,i}$ to the equations gives the following expression for a χ^2 function which will be minimized in the fitting process:

$$\chi^2 = \sum_{i=1}^{3} \left(\mathrm{w}_{1,i} \cdot \left(\tan(\theta_i) - \frac{\sqrt{(\mathrm{x}_b - (\rho_i \cos(\phi_i) + \mathrm{x}_{PA,i})^2 + (\mathrm{y}_b - (\rho_i \sin(\phi_i) + \mathrm{y}_{PA,i}))^2}}{\mathrm{z}_b - \mathrm{z}_{s,i}} \right)^2 + \mathrm{w}_{2,i} \cdot \left(\tan(\phi_i) - \frac{\mathrm{y}_b - \mathrm{y}_{PA,i}}{\mathrm{x}_b - \mathrm{x}_{PA,i}} \right)^2 \right) \tag{16}$$

The weighting factors w_1 and w_2 are given by the inverse of the uncertainties of the respective summands. These uncertainties are estimated using Gaussian error propagation.

The results of the fit for the stations 11, 21 and 23 for both χ^2 functions (Equation 11, respectively Equation 16) are shown in Figure 39. In this case, the GPS coordinates of the balloon are used as initial parameters for the fitting variables $\mathrm{x_b}$, $\mathrm{y_b}$ and $\mathrm{z_b}$. The additional initial parameters for the ϕ_i are calculated using the GPS coordinates together with Equation 15. The angles ϕ_i need to be able to account for values lying in the interval $[-\pi, \pi]$, and thus the `np.arctan2()` function is used [34].

Taking Figure 39, the result of the fit including the angles ϕ_i matches the expectation of the easting better than the fit based on the simpler χ^2 function. An exception is the last data point. However, for the northing and the altitude, the values of the more complex fit deviate further from the GPS track as the final values of the fit without auxiliary variables.

Although this fit provides more reasonable values for the x-coordinate compared to the simpler fit, it was not investigated further. This is based on the observation that the fit needs practically ideal input parameters to converge. The output of the fit provides a potential explanation for this behavior.

The output of `iminuit` of the more complex fit for the first data point is shown in Table 4. Again, ideal initial values are provided. It can be seen that the fit is only approximate due to a non-positive definite covariance matrix. According to [31] this can be a sign of an underdetermined system.

Tab. 4: Iminuit output for the fit with auxiliary variables. FCN is the last value of the χ^2 function considered by the minimization. EDM is the Estimated Distance to Minimum. NFCN describes the number of iterations. The other fields indicate if the fit was successful [31].

FCN = 0.07107 EDM = 1.28e-05 (Goal: 0.0002)		Nfcn = 235		
Valid Minimum		No Parameters at limit		
Below EDM threshold (goal × 10)		Below call limit		
Covariance	Hesse ok	APPROXIMATE	NOT pos. def.	FORCED

From the covariance correlation matrix (cf. Table 5) of the same data point, it can be seen that for example the correlation between the angle ϕ_0 and $\mathrm{y_b}$ is one. These properties are therefore fully correlated. This leads to the result that the system is indeed under-determined as one of those variables does not provide additional information. Thus, the fit does not converge for initial values that are not ideal.

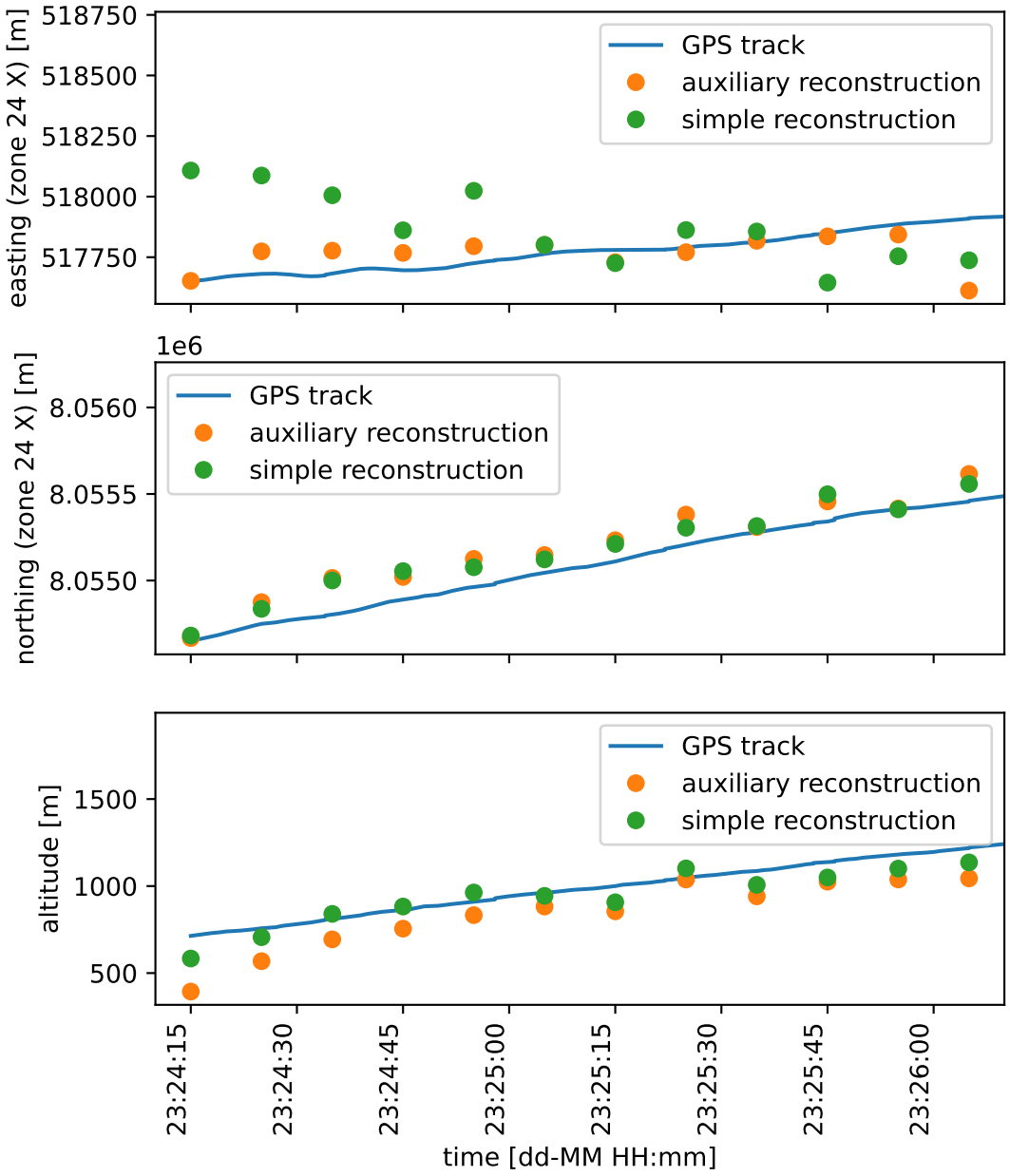


Fig. 39: Comparison of fit results from different χ^2 functions.

Tab. 5: Fit covariance correlation matrix

	x_b	y_b	z_b	ϕ_0	ϕ_1	ϕ_2
x_b	1	0.3	0	0.3	−0.4	0.7
y_b	0.3	1	0	1	−1	−0.5
z_b	0	0	1	0	0	0
ϕ_0	0.3	1	0	1	−0.9	−0.5
ϕ_1	−0.4	−1	0	−0.9	1	0.4
ϕ_2	0.7	−0.5	0	−0.5	0.4	1

This more complex fit model is expected to provide a more accurate fit result when additional information is available, such as the x- and y- coordinates of the surface interaction point or the angle in the xy-plane. As explained in section 2.2.3, each station has two helper strings in addition to the power string. The helper strings are also equipped with two Vpol antennas each. Using the time delay of the signal between the phased array and the antennas of the helper strings, the angle in the xy-plane can be estimated (phasing) under certain conditions. If the signal is pulsed, the time delay can be measured and the angle can be estimated. However, the signal emitted by the balloon is not pulsed but continuous. In case of a continuous signal, a well-defined angle can only be estimated if the wavelength is larger than the distance between the power string and the helper strings. For smaller wavelengths, several peaks with similar maximum power values are expected as described and observed in section 3.5.2. Unlike in the estimation of the zenith angle at the phased array, a general limitation of the angle in the xy-plane reducing the number of peaks to one is not possible. This approach is therefore not applicable for the balloon reconstruction ($\lambda = 0.4\,\mathrm{m} <<$ distance power string to helper strings, cf. Figure 2). However, it might be reasonable for the position reconstruction of other sources.

6 Conclusion

The goal of this thesis was to verify if it is possible to reconstruct the position of a radiosonde emitting a continuous signal using multiple in-ice phased array antennas of the RNO-G detector. Based on the different sub-steps of the reconstruction, further attention was paid to possible inaccuracies in the calibration of the phased array.

The second balloon flight launched on 03.08.2022 was investigated in this thesis.

In order to determine the angle at which the continuous wave impinges on the phased array antennas, the delay-and-sum beamforming technique was used (cf. section 3), also referred to as phasing. The process was optimized, e.g. by filtering the traces. The time evolution of the zenith angles at the phased array obtained by the beamforming technique was compared to the expected values calculated trigonometrically respectively using a ray tracing technique. It was shown that the shape of the evolution match each other. However, it could be observed that the absolute values for some stations deviate more clearly from the expectation than the values at other stations. This deviation might be due to inaccuracies in the calibration of the cable delays.

Furthermore, phasing was performed for different channel combinations. It was observed, that the resulting peaks for one event are positioned at similar zenith angles for a station where the absolute angle values obtained match the expectation. For a station where the reconstructed angle values deviated more clearly from the expectation, a shift in the peak positions was observed for different channel combinations. However, the analysis of more data is necessary to ensure reproducibility of the issue. Based on these observations, a method was proposed to detect and quantify deviations in the calibration of cable delays, or deviations in the distances between antennas (cf. section 3).

An alternative method for the detection of inaccuracies regarding the spacing of the antennas was proposed in section 4. It was based on determining the refractive index at which the obtained angle value matches the expectation best. It was expected that the obtained refractive index matches the known refractive index at the depth of the phased array. It was observed that the results for two stations met the expectation, whereas the analysis for three other stations led to a deviation from the expectation. Again, the analysis should be extended to more data, but based on the current result, it would be reasonable to carefully revise the spacing between the antennas of the stations in question.

The reconstruction of the balloon position was realized by a χ^2 minimization (fit). A demonstration of Gaussian error propagation for the unweighted χ^2 function F showed that a small deviation of the angle at the phased array resulted in a large deviation in the value of F.

Several fits were carried out, taking different station combinations, respectively different numbers of stations into account. For all combinations, the y- and z-coordinates of the balloon were well reconstructed. The reconstruction of the x-coordinate, on the other hand, was subject to larger deviations. Considering a simpler two dimensional case, these deviations could be traced back to be a combined issue of the station arrangement and angular inaccuracies present in one of the observed stations.

Finally, the results of the χ^2 minimization were compared to the results of a more complex χ^2 function, which used auxiliary variables. The latter fit was only converging for initial values similar to the GPS coordinates of the balloon. However, this more complex fit could be used to reconstruct pulsed signals, using the phased array and the helper string additionally.

As could be seen in this thesis, the use of interference sources can provide important clues to potential issues in calibration and antenna spacing. A further analysis of the phased array data in comparison with the GPS positions of the radiosonde is therefore reasonable. Especially the consideration of further balloon flights is recommended. For this purpose, the code developed in the course of this thesis is made available (`https://github.com/CarolinKl/radiosonde_tracker`).

A Time Evolution of Zenith Angles

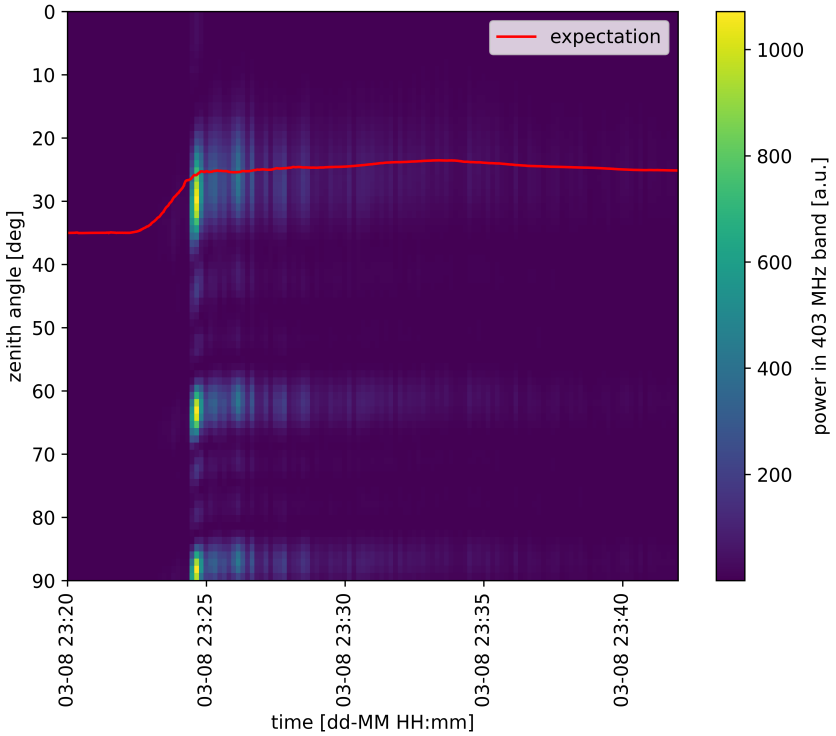


Fig. 40: Time evolution of zenith angles with trigonometric expectation of the zenith angle in-ice for station 11.

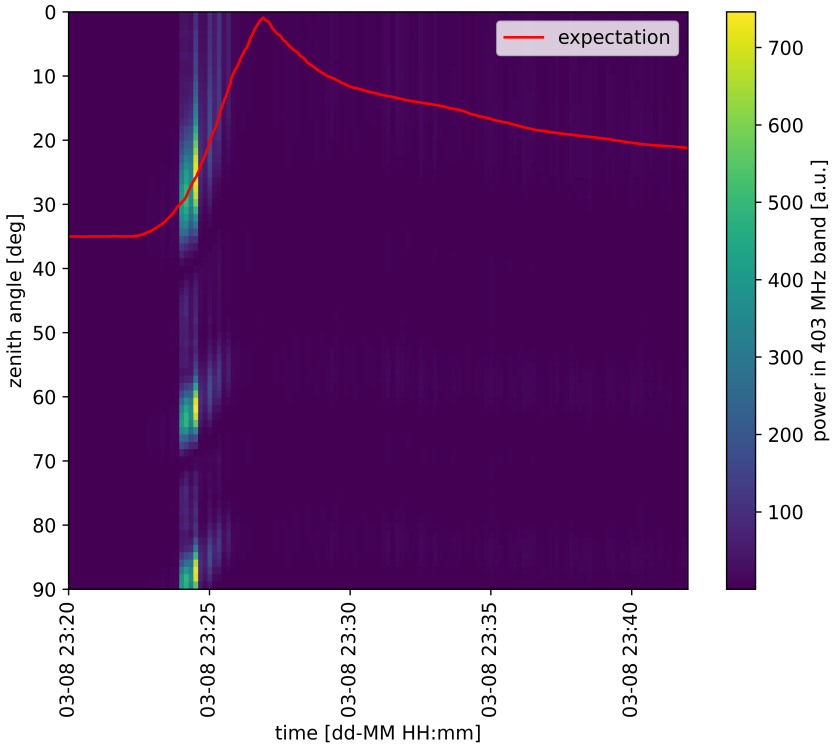


Fig. 41: Time evolution of zenith angles with trigonometric expectation of the zenith angle in-ice for station 22.

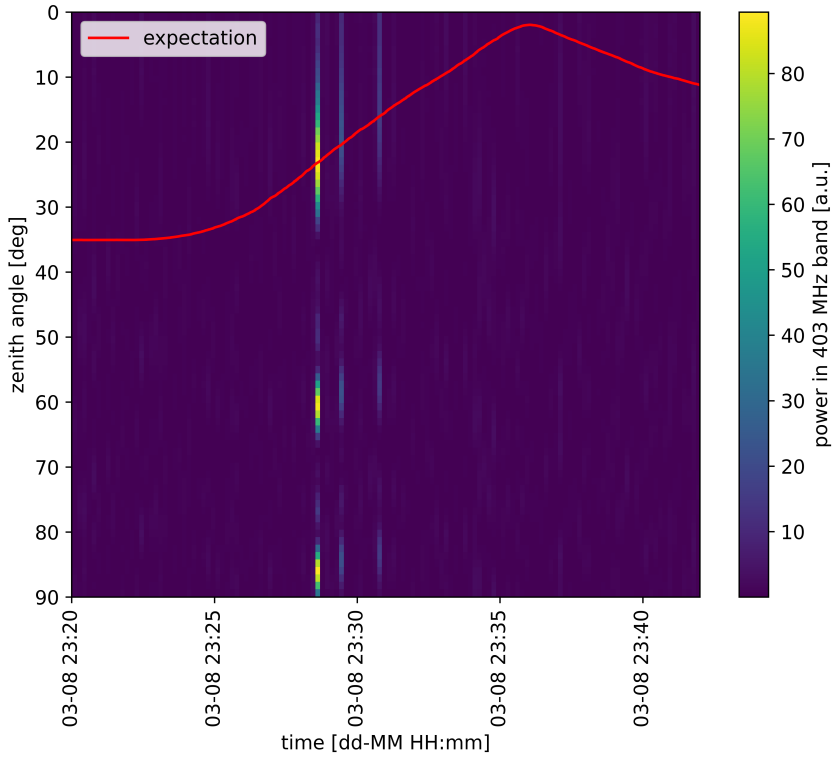


Fig. 42: Time evolution of zenith angles with trigonometric expectation of the zenith angle in-ice for station 24.

B Fit Parameters of Gaussian fits

The following values are not rounded, but are the values returned by `scipy.optimize.curve_fit`.

Tab. 6: Fit parameters Gaussian peak of station 11. Index denotes the position of the data point in the reconstruction fit.

index	a [a.u.]	m [deg]	sigma [deg]	d [a.u.]
0	13.344207093428818	32.3118878238519	3.868646385531159	1.0099649919939533
1	582.5194052561843	30.824855802022935	4.251787313504302	25.44079351051261
2	1019.557704418512	29.13459535661397	-4.451354659224496	44.43891370054234
3	529.5029940208814	27.724027044842877	4.535375184592349	26.91390258429215
4	235.35992158948557	28.102065985180683	4.749089111546608	6.818875626897931
5	292.17961286598114	26.69066704215538	4.698430163556344	12.994883234825254
6	310.98553938880787	26.765721395754838	5.0823902704736605	11.958419620584543
7	221.012769124063	26.079315006933598	-5.472812901336296	7.182045993816
8	234.21307038690637	27.05543384600013	4.959883272616233	9.56702998773016
9	174.25443138053373	25.855753927105887	5.007539216322532	6.4040357477348
10	290.267651376552	25.690003328091258	4.962008088785329	10.70619142451459

Tab. 7: Fit parameters Gaussian peak of station of station 21. Index denotes the position of the data point in the reconstruction fit.

index	a [a.u.]	m [deg]	sigma [deg]	d [a.u.]
0	60.25460927679276	14.355110202270978	7.266756295370744	2.283766709925484
1	33.7609231268961	15.215087333123384	6.664792512241961	2.9907891531323982
2	113.44802553504068	16.131212584583235	7.284440566611424	3.6977773771561733
3	125.95741525411663	16.159416054387982	7.099389504375617	8.962383747486914
4	68.6823968821176	16.36165334569937	7.087556460816478	4.4684438458637
5	76.89348286711372	16.932427459489737	6.939902517787666	7.4811857523413305
6	34.28305951471319	19.475535814892307	6.5459820695509014	2.1803110795858363
7	33.78957001924506	18.561470624444624	6.1988671363035905	4.366952072599139
8	13.961659214843875	19.886902794460884	5.280079697201079	3.310863208501239
9	28.183423451314958	22.470210846539462	5.91705599872236	1.63116143529684
10	63.46420634312597	20.374697972246825	6.194293366292333	7.635858202720257

Tab. 8: Fit parameters Gaussian peak of station of station 22. Index denotes the position of the data point in the reconstruction fit.

index	a [a.u.]	m [deg]	sigma [deg]	d [a.u.]
0	340.56397764756815	26.55442163076854	5.595790323765003	25.39047404253255
1	665.0433093586438	24.431249181368138	5.963253553476872	79.07558353279359
2	81.14121344608898	20.73546661224005	6.8975709308867446	12.289652093872453
3	65.60567596707092	19.256697474279683	7.055982780062836	-1.4640389185853895
4	198.32137499314007	18.210921042494366	7.980865026095799	-13.484562062230584
5	76.52771308148988	16.336732472374315	6.782743764372843	1.568395806843021
6	119.98727460596746	14.166165025664199	6.49557734053289	6.280388409875783
7	25.892594983050653	12.895766107505224	6.772262269080993	0.09913280994689808
8	55.26259289110065	12.690936055857547	6.04472788627699	2.1276172524513783
9	10.505602880203211	13.023575641059109	6.333117178404821	0.4423954507071408
10	9.421603549252666	12.795451541554721	6.330537164999962	0.7050260001108332

Tab. 9: Fit parameters Gaussian peak of station of station 23. Index denotes the position of the data point in the reconstruction fit.

index	a [a.u.]	m [deg]	sigma [deg]	d [a.u.]
0	57.16188087720409	33.43624235718461	3.847897645081435	5.434589531394532
1	51.94638799093995	32.84604330766699	4.105326623674072	2.850907507999778
2	88.84555062899774	31.649386328530518	4.05309831382051	6.850763274981897
3	88.55862647336748	31.252597038210983	4.368904015685591	3.3850910196336175
4	105.40156107694924	30.495999428501023	4.27739444610914	8.924459321784989
5	64.81061935469558	30.60933445760475	4.021478539892963	5.199259630327171
6	101.24498192589006	30.621338243644594	4.0861397970579265	6.552121869671608
7	115.27809183479289	28.396421518542223	4.731820562560068	6.460418729479157
8	79.06501883494016	29.171415002527606	4.608949505148535	4.332449283782613
9	78.97654570498153	28.110917310161767	4.504684047536025	5.035975961684975
10	55.483487220306515	27.94306515526636	4.84361803439153	2.845829837257606

References

[1] J. Aguilar, P. Allison, J. Beatty, *et al.*, "Design and sensitivity of the radio neutrino observatory in greenland (rno-g)," *Journal of Instrumentation*, vol. 16, p. P03025, mar 2021.

[2] M. Shupe, D. Turner, V. P. Walden, *et al.*, "High and dry: New observations of tropospheric and cloud properties above the greenland ice sheet," *Bulletin of the American Meteorological Society*, vol. 94, no. 2, pp. 169–186, 2013.

[3] P. Allison, S. Archambault, R. Bard, *et al.*, "Design and performance of an interferometric trigger array for radio detection of high-energy neutrinos," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 930, pp. 112–125, 2019.

[4] C. Sutton, *Spaceship Neutrino*. Cambridge University Press, 1992.

[5] N. Schmitz, *Neutrinophysik*. Vieweg+Teubner Verlag, 1997.

[6] M. Aartsen, M. Ackermann, J. Adams, *et al.*, "Neutrino astronomy with the next generation icecube neutrino observatory," 2019.

[7] E. Paudel, A. Coleman, and F. Schroeder, "Parametrization of the Relative Amplitude of Geomagnetic and Askaryan Radio Emission from Cosmic-Ray Air Showers using CORSIKA/CoREAS Simulations (ICRC2021)," jul 2021.

[8] S. Barwick and C. Glaser, "Radio detection of high energy neutrinos in ice," aug 2022.

[9] C. Grupen, *Astroparticle Physics*. Springer International Publishing, 2020.

[10] F. Schröder, "Radio detection of cosmic-ray air showers and high-energy neutrinos," *Progress in Particle and Nuclear Physics*, vol. 93, pp. 1–68, 2017.

[11] B. Saleh and M. Teich, *Optik und Photonik*. John Wiley & Sons, 2020.

[12] International Telecommunication Union, "The Radio Refractive Index: Its Formula and Refractivity Data," Tech. Rep. Recommendation ITU-R P.453-14, 2019.

[13] S. Barwick, E. Berg, D. Besson, *et al.*, "Observation of classically 'forbidden' electromagnetic wave propagation and implications for neutrino detection," *JCAP07(2018)055*, vol. 2018, pp. 055–055, jul 2018.

[14] C. Deaconu, A. Vieregg, S. Wissel, *et al.*, "Measurements and modeling of near-surface radio propagation in glacial ice and implications for neutrino experiments," *Phys. Rev. D 98, 043010 (2018)*, vol. 98, p. 043010, aug 2018.

[15] NuRadio Group, "NuRadioMC Documentation: Ice and Attenuation Models." `https://nu-radio.github.io/NuRadioMC/NuRadioMC/pages/Manuals/icemodels.html#id1`, Accessed: 10.08.2023.

[16] S. Hallmann, "New physics and ultra-high-energy neutrinos." `https://indico.e5.physik.tu-dortmund.de/event/1004/attachments/803/1624/Lecture_UHEnu_newPhysics.pdf`, 2021. Accessed: 01.06.2023.

[17] I. Plaisier, *Reconstructing the Arrival Direction of Cosmic Neutrinos with the Radio Neutrino Observatory Greenland (RNO-G)*. PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2022.

[18] G. Ohring, "a most surprising discovery," *Bulletin of the American Meteorological Society*, vol. 45, no. 1, pp. 12–14, 1964.

[19] International Telecommunication Union, "Technical characteristics and performance criteria for systems in the meteorological aids service in the 403 MHz and 1 680 MHz bands." https://www.itu.int/dms_pubrec/itu-r/rec/rs/R-REC-RS.1165-3-201812-I!!PDF-E.pdf, 2018. Accessed: 01.08.2023.

[20] A. Kräuchi, R. Philipona, G. Romanens, *et al.*, "Controlled weather balloon ascents and descents for atmospheric research and climate monitoring," *Atmospheric measurement techniques*, vol. 9, no. 3, pp. 929–938, 2016.

[21] Vaisala Corporation, "Vaisala Radiosonde RS41 Datasheet." https://www.vaisala.com/sites/default/files/documents/WEA-MET-RS41-Datasheet-B211321EN.pdf. Accessed: 18.7.2023.

[22] Pictures taken by J. Henrichs.

[23] J. Manchuk and C. Deutsch, "Conversion of latitude and longitude to utm coordinates," *Paper 410, CCG Annual Report*, vol. 11, 2009.

[24] B. Oeyen. https://github.com/RNO-G/analysis-tools/blob/main/rnog_analysis_tools/coordinate_system/coordinate_origins.json. Accessed: 25.05.2023.

[25] "Giovanni measurement definitions: Geopotential height." https://disc.gsfc.nasa.gov/information/glossary/59a8444857c4c46c50e8dc6e. Accessed: 19.07.2023.

[26] Y. Deng, A. Ridley, and W. Wang, "Effect of the altitudinal variation of the gravitational acceleration on the thermosphere simulation," *Journal of Geophysical Research: Space Physics*, vol. 113, no. A9, 2008.

[27] "RNO-G Station Map." https://rno-g.github.io/station-map/. Accessed: 01.06.2023.

[28] C. Glaser, D. García-Fernández, A. Nelles, *et al.*, "NuRadioMC: simulating the radio emission of neutrinos from interaction to detector," *The European Physical Journal C*, vol. 80, jan 2020.

[29] R. Krause, *Antenna development and calibration for measurements of radio emission from extensive air showers at the Pierre Auger Observatory*. Dissertation, RWTH Aachen University, Aachen, 2018.

[30] B. Oeyen, J. Aguilar, P. Allison, *et al.*, "The interplay of ice-firn model and station calibration in RNO-G," *Proceedings of the 38th International Cosmic Ray Conference (ICRC2023)*, July 2023. https://pos.sissa.it/444/1042/pdf. Accessed: 15.08.2023.

[31] F. James, "Minuit," *CERN Program Library Long Writeup D*, vol. 506, p. 1993, 1994.

[32] H. Dembinski, P. Ongmongkolkul, C. Deil, *et al.*, "iminuit: Jupyter-friendly python interface for c++ minuit2," *Astrophysics Source Code Library*, pp. ascl–2108, 2021.

[33] R. Andrae, T. Schulze-Hartung, and P. Melchior, "Dos and don'ts of reduced chi-squared," Dec. 2010. https://arxiv.org/pdf/1012.3754.pdf. Accessed: 16.08.2023.

[34] NumPy Developers, "numpy.arctan2." https://numpy.org/doc/stable/reference/generated/numpy.arctan2.html. Accessed: 29.06.2023.

Erklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbst verfasst habe. Weiterhin habe ich nur die angegebenen Quellen und Hilfsmittel verwendet.

Erlangen, 21.08.2023 Carolin Klein