Documentation of the
DWD Data Assimilation System

Main Authors:

Andreas Rhodin, Heiner Lange, Roland Potthast, Tijana Janjic-Pfander

With Contributions By:

Harald Anlauf, Heinz-Werner Bitzer, Alexander Creß, Robin Faulwetter,
Christina Köpken-Watts, Martin Lange, Gerhard Paul, Hendrik Reich,
Christoph Schraff, Klaus Stephan, Olaf Stille, Annika Schomburg,
Michael Bender, Rory Gray, Africa Perianez, Jason Ambadan,
Eric Lange, Stefan Declair, Detlef Pingel, Oliver Schmid,
Nicolas Fischer

March 20, 2015

Todo list

About this Documentation

This *Documentation of the DWD Data Assimilation Systems* has two aims:

1. To provide new users an introduction to the scientific background of Data Assimilation (DA) and its implementation at the DWD

2. To serve as a work of reference for the developers and users of the respective subsystems.

The documentation is organized as follows:

Part **??** **Data Assimilation and the Cycled NWP System** displays the basic ideas of data assimilation and how it is organized at DWD. This part of the documentation serves as an introduction to the further parts that are linked in there.

Part **??** **Scientific Documentation** describes the algorithms of the different data assimilation methods.

Part **??** **User Guide** shows the user how to install and execute the system's components. Herein, the namelists are described and guidance is provided how to run cycled data assimilation.

Part **??** **Implementation** describes the codes and modules.

The sequence of these Parts is such that a continous reading of the document should be consistently possible. The Parts contain reading guidelines for themselves.

The typical reader will probably skip back and forth through the document – in that case, *forward-links* from the scientific part to the implementation and user guide are provided, and occasionally *backward-links*, especially in the case of namelist variables in the user guide that act as switches for certain algorithmic features.

It is important to acknowledge that this documentation is always *work-in-progress*, because the features and algorithms of the DWD DA System are always in development.

If the reader has the feeling that something is missing or badly/wrongly documented, he/she should not hesitate to contact the authors or to actively contribute to this documentation, following the writing guidelines (Chapter **??**) in the Appendix.

Contents

Chapter 1

Observation Operators

1.1 Observation operators for global models

1.1.1 GNSS Slant Total Delay Operator

Ground based GNSS[1] makes use of GNSS data collected by large GNSS networks. GNSS processing algorithms provide the optical path length for each individual satellite-receiver link, i. e. information about the atmospheric state integrated along the signal path. The slant total delay (STD) is the atmospheric contribution to the optical path length. These quantities are called delays or execess path lengths as the atmosphere delays the signal as compared to undisturbed propagation in vacuum. Zenith total delays (ZTDs) are the correspondig delays in zenith direction and are equivalent to STDs with elevations of 90°. However, ZTDs are usually processed in a different way involving all available observations within a given period which are combined to a hypothetical observation in zenith direction.

ZTD data can be obtained from several network providers or GNSS processing centers in the world. The data are usually available in near real-time or as real-time stream. Near real-time data are processed in temporal batches, e. g. one hour, and provided with a certain delay, e. g. 75 minutes. ZTD data for Europe can be obtained from E-GVAP which is an association of several European GNSS processing centers and weather services. Operational STD data can currently[2] obtained only from the GFZ in Potsdam. The GFZ provides near real-time STD data with a delay of about 75 minutes and a focus on Germany.

The observed STD is an integral of the atmospheric refractive index along the signal path from the GNSS satellite to the GNSS receiver at the ground. The refractive index is related to the pressure, temperature and humidity fields and the length of the signal path inside the model ranges from the vertical extension of the model (ZTD, 90° elevation) up to several hundred km or even more than 1000 km, depending on the elevation and the height of the model top layer. STD observations are therefore neither localized nor related to one specific meteorological quantity.

An other issue is the signal path inside the atmosphre which is not known in advance but must be estimated from the background model using a raytracer. GNSS data sets provide the coordinates of the GNSS satellite and of the GNSS station, the signal path in between depends on the specific atmospheric state in the vicinity of the connecting line and can be estimated by solving Fermat's principile for a given model state. The resulting bended signal path has a considerable effect on the path delay and cannot be neglected for observations with elevations smaller than $\sim 30°$.

Observation Operator

The STD observation operator consits of two parts. In the first part a connecting line between the station and the satellite is defined and a set of supporting points along this line is chosen. Using the coordinates of these supporting points the interface to the assimilation system collects the required meteorological data and provides them to the second part of the operator. As the GNSS signal can propagate through large parts of the model access to grid nodes with quite large distances to the station is often required.

---

[1]The generic term GNSS (Global Navigation Satellite Systems) refers to all existing navigation satellite systems, i. e. GPS, GLONASS, Compass, Galileo, ...

[2]Summer 2014

The second part of the STD operator calls a raytracer in order to estimate the bended signal path and integrates along this path to obtain the STD.

The basic steps of the STD operator are:

1. Define a connecting line between the GNSS station and the GNSS satellite and a set of supporting points on this line which will later be used for numerical integration. The point density is scaled with the pressure profile in order to obtain point distances which decrease with increasing height.

2. Collect the temperature, pressure and humidity data on the grid nodes surrounding the supporting poins.

3. Compute the atmospheric refractivity on these nodes using the Smith & Weintraub formula.

4. Interpolate the refractivities at the supporting points.

5. Call the raytracer to estimate the signal path in the atmosphere, i.e. the deviations from the connecting line. The raytracer is implemented as a iterative Newton algorithm which minimizes the optical path length in the atmosphere (Fermat's principle). The iteration is started with the connecting line and makes small modifications in order to approximate the "true" bended signal path. Within this step the refractivities need to be interpolated on the latest approximation of the signal path.

6. The STD is the line integral through the refractivity field along the signal path. This is obtained by numerical integration.

7. Ionospheric effects are neglected.

The STD operater can also be used to assimilate ZTD data as the ZTD is a specific case of a STD with an elevation of 90°. However, due to different GNSS procesing strategies the ZTD errors should be smaller than the corresponding STD errors. The same STD operator code is used in GME/ICON and COSMO, only the interfaces are different.

Incomplete: GNSS Slant Total Delay operator

Chapter 2

Namelist Groups

2.1 /STD_OBS/ (GNSS slant delay operator)

Empty: GNSS Slant Total Delay operator namelist

This namelist is defined in module STD_operator :

```
!----------------------------
! namelist STD_OBS
!----------------------------
logical           :: read_ascii  = .false.! read ascii file
integer ,parameter :: mo          = 12     ! max number of observation
                                           ! input files
integer ,parameter :: mf          = 256    ! max length of file names
character (len=10) :: RefTimeStr = ''      ! Reference time (analysis)
                                           ! yyyymmddhh
integer           :: STDperiod   =  0      ! STD period, minutes
integer           :: NStepVertMod = 60     ! Number of vertical points
                                           ! inside the model
integer           :: NStepVertTop = 80     ! Number of vertical points
                                           ! above the model
real (wp)         :: Hmax    = 150000.     ! Maximum hight for STD
                                           ! integration in m
real              :: HScaleP = 6500.       ! air pressure scale hight
                        ! (for scaling the point density along the slant)
real              :: HScaleP2 = 25000.     ! air pressure scale hight
real (wp)         :: k1 = 77.60_wp         ! refractivity coefficients
real (wp)         :: k2 = 70.40_wp         ! refractivity coefficients
real (wp)         :: k3 = 3.739E5_wp       ! refractivity coefficients
character(len=mf)  :: SlantPath = ''        ! Path to slant data
character(len=mf)  :: std_obs_file (mo) ='' ! STD observation input files
character(len=mf)  :: HORIFile = 'stako_gps_SNX_G_12_HORI'
                                           ! GNSS station file
character(len=mf)  :: DomesFile = 'smark_domes_SNX_RTT'
                                           ! GNSS station locations
```

Parameter	Default	Description
NStepVertMod	60	Number of supporting points used for vertical integration inside the model, should be $\geq$ the number of vertical grid nodes of the model. The number of supporting points on the signal path is scaled with the elevation: Nmod = NStepVertMod / $(2\sin\varepsilon)$.
NStepVertTop	80	Number of supporting points used for vertical integration above the model top. NStepVertTop can usually be smaller than NStepVertMod as the contributions from the upper atmosphere is small. Like NStepVertMod this is scaled with the elevation: Nup = NStepVertTop / $(2\sin\varepsilon)$. The total number of supporting points on the signal path is Ntot = 1 + (NStepVertMod + NStepVertTop) / $(2\sin\varepsilon)$.
Hmax	150000	The STD is integrated from the GNSS station up to Hmax in m. $100000 \leq$ Hmax ≤ 200000 (100 km … 200 km)
HScaleP	6500	Scale hight in m of a hypothetical pressure profile. This profile is used to scale the density of supporting points with height. Smaller numbers lead to more points in the lower atmosphere. HScaleP is used for the signal path inside the model.
HScaleP2	25000	Like HScaleP but used above the model top.
k_1	77.60	Refractivity coefficient k_1 [K hPa-1] used in the Smith & Weintraub formula or the Thayer formular, default according to Bevis, 1994.
k_2	70.40	Refractivity coefficient k_2 [K hPa-1], default according to Bevis, 1994.
k_3	$3.739 \cdot 10^5$	Refractivity coefficient k_3 [K^2 hPa-1], default according to Bevis, 1994.

The remaining namelist parameters are closely related to the preliminary data format and will not be described here.

Refractivity coefficients according to differnt authors and their impact on a hypothetical ZTD (last column):

Quelle	k_1 [K hPa^{-1}]	k_2 [K hPa^{-1}]	k_3 [K^2 hPa^{-1}]	ZTD [m]
Smith & Weintraub	77.607	71.60	$3.747 \cdot 10^5$	2.417459855
Thayer	77.604	64.79	$3.776 \cdot 10^5$	2.417989384
Hasagawa	77.60	69.40	$3.701 \cdot 10^5$	2.414152786
Bevis	77.60	70.40	$3.739 \cdot 10^5$	2.416587827
Foelsche	77.650	65.99	$3.777 \cdot 10^5$	2.419539541
Rueger	77.689	71.2952	$3.75463 \cdot 10^5$	2.420170997

The refractivity coefficients suggested by Bevis seem to provide the best results and have been chosen as default.

Add references

Chapter 3

Source Code Reference

3.1 Observations

> **Empty: GNSS Slant Total Delay operator: Theory**

Signal Delays

The optical path length L is defined by the integral

$$L = \int_S n(s)\, ds \tag{3.1}$$

S beeing the true bended signal path in the atmosphere and $n(s)$ beingt the refraction index along this path. The optical path difference ΔL as defined by

$$\Delta L = \int_S n(s)\, ds - \int_G ds \tag{3.2}$$

is the difference between L and the undisturbed signal propagation in vacuum ($n = 1$). The signal path in vacuum is a straight line G and $\int_G ds$ is the geometric distance between transmitter and receiver. After some algebra two contributions to ΔL can be identified:

$$\Delta L = \int_S n(s)\, ds + \int_S ds - \int_S ds - \int_G ds \tag{3.3}$$

$$= \int_S (n(s) - 1)\, ds + \int_S ds - \int_G ds \tag{3.4}$$

$$= \int_S (n(s) - 1)\, ds + (S - G) \tag{3.5}$$

The first term $\int_S (n(s) - 1)\, ds$ describes the delay due to the delayd signal propagation in matter ($v < c$) and the second term describes the extra travel time due to the increased path length of the bended signal path (S-G). The latter is referred to as the *geometric delay*. The speed of light enters the equations implicitly as the refraction index $n = \frac{c}{v}$, v beeing the speed of light in matter, c beeing the vacuum speed of light. The term $n - 1$ is usually rather small, especially in gases. Therefore, the refractivity N is defined:

$$N = 10^6 \cdot (n - 1) \tag{3.6}$$

leading to

$$\Delta L = STD = 10^{-6} \cdot \int_S N\, ds + (S - G) \tag{3.7}$$

ΔL is referred to as the *slant total delay* (STD). The term $S - G$ is much smaller than the first term and is often neglected, leading to the well known approximation

$$STD \approx 10^{-6} \cdot \int_S N\, ds \tag{3.8}$$

This approximation is not used by the STD observation operator. At low elevations the STD is about 20 m – 25 m and the geometric delay is about 15 cm – 25 cm, i.e. approx. 1 % of the total delay. In zenith direction the STD (= ZTD) is about 2.4 m and the geometric delay is almost zero.

The STD provided by GNSS data processing is a quantity related to the neutral atmosphere [1], i. e. ionospheric effects were separated before the STD was estimated. As the STD observation operator must provide the same physical quantity the integral 1.7 has to deal only with the subpath S inside the neutral atmosphere. This is not necessarily true for the raytracing. The free charges in the ionosphere have a strong impact on the signal path. i. e. on the point where the signal enters the neutral atmosphere. However, this effect is small and is neglected by the STD operator.

Raytracing

According to Fermat's principle each electromagnetic wave propagates along a path S which minimizes the optical path length:

$$\int_S n(s)\, ds \;\Rightarrow\; \text{min.} \tag{3.9}$$

Assuming that the field of the refracion index n is known, this equation can be used to estimate the bended signal path in the atmosphere. The variational problem is best solved in a Cartesian system with the x axis pointing from the receiver to the satellite. The y and z coordinates of the signal path become functions of x which describe the deviations from a straight line:

$$\int_a^b n(x,y,z)\sqrt{1 + y'(x)^2 + z'(x)^2}\, dx = \int_a^b F(x, y(x), y'(x), z(x), z'(x))\, dx \;\Rightarrow\; \text{min.}$$

The variational problem can be translated to a system of partial differential equations using the Euler-Lagrange equations:

$$\frac{\partial F}{\partial y} - \frac{d}{dx}\frac{\partial F}{\partial y'} = 0 \;\Rightarrow\; \frac{\partial F}{\partial y} - \frac{\partial^2 F}{\partial x \partial y'} - \frac{\partial^2 F}{\partial y \partial y'} - \frac{\partial^2 F}{\partial y'^2} = 0$$

$$\frac{\partial F}{\partial z} - \frac{d}{dx}\frac{\partial F}{\partial z'} = 0 \;\Rightarrow\; \frac{\partial F}{\partial z} - \frac{\partial^2 F}{\partial x \partial z'} - \frac{\partial^2 F}{\partial z \partial z'} - \frac{\partial^2 F}{\partial z'^2} = 0 \tag{3.10}$$

This leads to two coupled ordinary differential equations for the functions $y(x)$ and $z(x)$

$$y'' = \left(\frac{n_y}{n} - \frac{n_x}{n}y'\right)(1 + y'^2 + z'^2)$$

$$z'' = \left(\frac{n_z}{n} - \frac{n_x}{n}z'\right)(1 + y'^2 + z'^2) \tag{3.11}$$

with

$$n_x = \frac{\partial\, n(x,y,z)}{\partial x} \;,\; n_y = \frac{\partial\, n(x,y,z)}{\partial y} \;,\; n_z = \frac{\partial\, n(x,y,z)}{\partial z} \tag{3.12}$$

---

[1] In the GNSS literature the term "troposphere" is very often used as a sysnonym for the neutral atmosphere which reaches up to ~ 100 km. Regarding the propagation of radio waves the atmosphere is divided into the neutral atmosphere and the ionosphere and everything below the ionosphere is regarded as "troposphere".

The equations 3.11 have to be solved with the boundary conditions

$$y(a) = z(a) = y(b) = z(b) = 0 \tag{3.13}$$

which guarantee that the signal propagates from the transmitter at $x = b$ to the receiver at $x = a$.

For the numerical solution a set of supporting points on the x axis needs to be defined:

$$x_i , \quad i = 1, \ldots, N , \quad \text{with } x_1 = a , \quad x_N = b$$

Using these x_i the differential equations could be approximated by finite differences but a more general solution can be found by connecting neigbored points by polynomials and differentiating the polynomials. The Lagrange polynomials $L(x)$ were chosen to connect sets of three neighbored points $(x_{i-1},\ x_i,\ x_{i+1})$:

$$L(x) = \sum_{k=i-1}^{k=i+1} f_k \cdot l_k(x) , \quad l_k(x) = \prod_{\substack{m=i-1 \\ m \neq k}}^{i+1} \frac{(x - x_m)}{(x_k - x_m)}$$

Here f_k is the value of the function at the node x_k, i. e. $f_k = y_k$ or $f_k = z_k$ and the l_k are the Lagrange basis polynomials. After replacing the functions $y(x)$ and $z(x)$ with the corresponding Lagrange polynomials and their derivatives the following system of algebraic equations is obtained:

$$\sum_{k=i-1}^{k=i+1} y_k l_k''(x_i) = \left(\frac{n_y}{n} - \frac{n_x}{n} \sum_{k=i-1}^{k=i+1} y_k l_k'(x_i) \right) \cdot$$
$$\left[1 + \left(\sum_{k=i-1}^{k=i+1} y_k l_k'(x_i) \right)^2 + \left(\sum_{k=i-1}^{k=i+1} z_k l_k'(x_i) \right)^2 \right] \tag{3.14}$$

$$\sum_{k=i-1}^{k=i+1} z_k l_k''(x_i) = \left(\frac{n_z}{n} - \frac{n_x}{n} \sum_{k=i-1}^{k=i+1} z_k l_k'(x_i) \right) \cdot$$
$$\left[1 + \left(\sum_{k=i-1}^{k=i+1} y_k l_k'(x_i) \right)^2 + \left(\sum_{k=i-1}^{k=i+1} z_k l_k'(x_i) \right)^2 \right] \tag{3.15}$$

The $2(N - 2)$ non-linear equations with $2(N - 2)$ unknowns can now be solved in order to obtain the y_i and z_i which define the functions $y(x)$ and $z(x)$ at the supporting points $x_i,\ i = 2, \ldots, N - 1$. The values $y_1 = y_N = 0$ and $z_1 = z_N = 0$ are defined by the boundary conditions and need not to be estimated. The derivatives of the Lagrange basis polynomials l_k are required only at the supporting points x_i and are simple combinations of the point distances Δx_i.

The equations XX can be rearenged and written in vector notation as $\boldsymbol{G}(\boldsymbol{x}) = 0$, where the vector function

$$\boldsymbol{G} = (g_{1,2}, g_{2,2}, \ldots, g_{1,i}, g_{2,i}, \ldots, g_{1,N-1}, g_{2,N-1})$$

consists of the functions $g_1(x_i) = g_{1,i}$ defined by equ. 3.14 and the functions $g_2(x_i) = g_{2,i}$ defined by equ. 3.15 evaluated at each supporting point x_i, $i = 2, \ldots, N - 1$. The vector $\boldsymbol{x}$ of unknown is given by

$$\boldsymbol{x} = (y_2, z_2, \ldots, y_i, z_i, \ldots, y_{N-1}, z_{N-1})$$

A system of non-linear equatins $\boldsymbol{G}(\boldsymbol{x}) = 0$ can be solved with the Newton algorithm which finds the roots of the equations iteratively:

$$\boldsymbol{x}^{k+1} = \boldsymbol{x}^k - \left(\nabla \boldsymbol{G}(\boldsymbol{x}^k)\right)^{-1} \cdot \boldsymbol{G}(\boldsymbol{x}^k)$$

In this form the inverse of the Jacobian $J = \nabla \boldsymbol{G}$ is required. In most cases it is more efficient to solve a system of linear equations which provides the corrections $\Delta \boldsymbol{x}^k$ to the latest estimate of the root $\boldsymbol{x}^k$:

$$\Delta \boldsymbol{x}^k \cdot \nabla \boldsymbol{G}(\boldsymbol{x}^k) + \boldsymbol{G}(\boldsymbol{x}^k) = 0 \quad \Rightarrow \quad \boldsymbol{x}^{k+1} = \boldsymbol{x}^k + \Delta \boldsymbol{x}^k \qquad (3.16)$$

The iteration continues until the norm $\left|\Delta \boldsymbol{x}^k\right|$ is sufficiently small and $\boldsymbol{x}^k$ is a good estimate of the root.

The Newton algorithm requires the Jacobian of $\boldsymbol{G}$, i. e. the partial derivatives of $\boldsymbol{G}$ with respect to the $2(N - 2)$ unknowns y_i and z_i. In equ. 3.14/3.15 it was decided to connect 3 neigbored supporting points by the interpolating Lagrange polynomial and consequently each single function g_i depend on only 6 variables:

$$g_i(\boldsymbol{x}) = g_i\left(y_{i-1}, z_{i-1}, y_i, z_i, y_{i+1}, z_{i+1}\right)$$

This leads to a Jacobian with only 6 non-zero entries per row. The consecutive functions $g_{1,i}$ and $g_{2,i}$ always depend on the same set of unknowns and the Jacobian is a band matrix of band width 7.

The band width depends on the number of supporting points which are connected by the interpolating Polynom. If r points are connected the band width is $2r + 1$. In case of the Lagrange polynomials $r = 3$ is the minimum number of points and 7 is the minimum band width. Chosing a larger r should lead to more accurate and stable results but increases the number of partial derivatives which need to be computed and also the computing time to solve a system of linear equation with a larger band width. It seems that $r = 3$ is entirely sufficient.

The system of linear equation defined by equ. 3.16 can be solved by any standard technique but there are especially efficient algorithms for solving band matrices.

The Newton algorithm does in general not provide stable solutions and its results can depend to a large degree on the start vector $\boldsymbol{x}_0$ of the iteration. Fortunately, the deviations from a linear signal propagation are rather small, i. e. several 100 m over distances of several 100 km, and a straight line ($\boldsymbol{x}_0 = 0$) is a good starting point. It turned out that even a rather simple implementation of a Newton algorithm leads to stable and fast results. Usually, $1 - 3$ iterations are sufficient to obtain accurate results.

The vector $\boldsymbol{x}$ provided by the Newton algorithm is the final solution for the bended signal path in the atmosphere. The supporting points x_i and the y_i and z_i from $\boldsymbol{x}$ define the bended signal path (x_i, y_i, z_i) which can in the next step be used to estimate the STD.
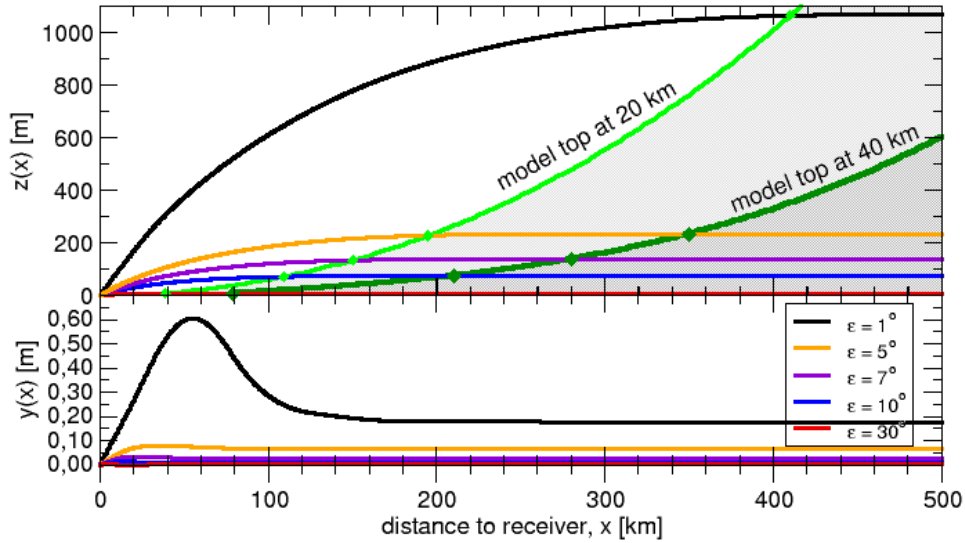
Figure 3.1: Raybendeing

The signal delay

After the raytracer provided an estimate of the signal path S in the atmosphere the signal delay can be computed using equ. 3.2. The second part $\int_G ds$ is the geometric distance between satellite and receiver which is known from the corresponding coordinates and the first part $\int_S n(s)\, ds$ was already computed within the last step of the raytracer:

$$STD = \int_a^b n(x,y,z)\sqrt{1 + y'(x)^2 + z'(x)^2}\, dx - \int_G ds$$

The derivatives of the discrete functions $y(x)$ and $z(x)$ are obtained by computing the derivatives of the corresponding Lagrange polynomials.

Reference systems, coordinates

geographische Koordinaten, ECEF, lokales Horizontsystem, Slant-System, ...

Transformations

Transformations between ellipsoidal and cartesian coordinates   The ellipsoidal coordinates longitude λ, latitude β and height h can be transformed into ECEF kartesian coordinates X, Y and Z and vice versa. It is assumed that the origin of the ECEF cartesian system, i.e. the center of Earth, is also the center of the ellipsoid.

1) Transformations from ellipsoidal to cartesian coordinates   Querkrmmungsradius des Ellipsoids:

$$N = \frac{a}{\sqrt{1 - e^2 \cdot \sin^2 \beta}} = \frac{a}{\sqrt{1 - f(2 - f)\sin^2 \beta}} = \frac{a^2}{\sqrt{a^2 + \cos^2 \beta + b^2 \sin^2 \beta}} \qquad (3.17)$$
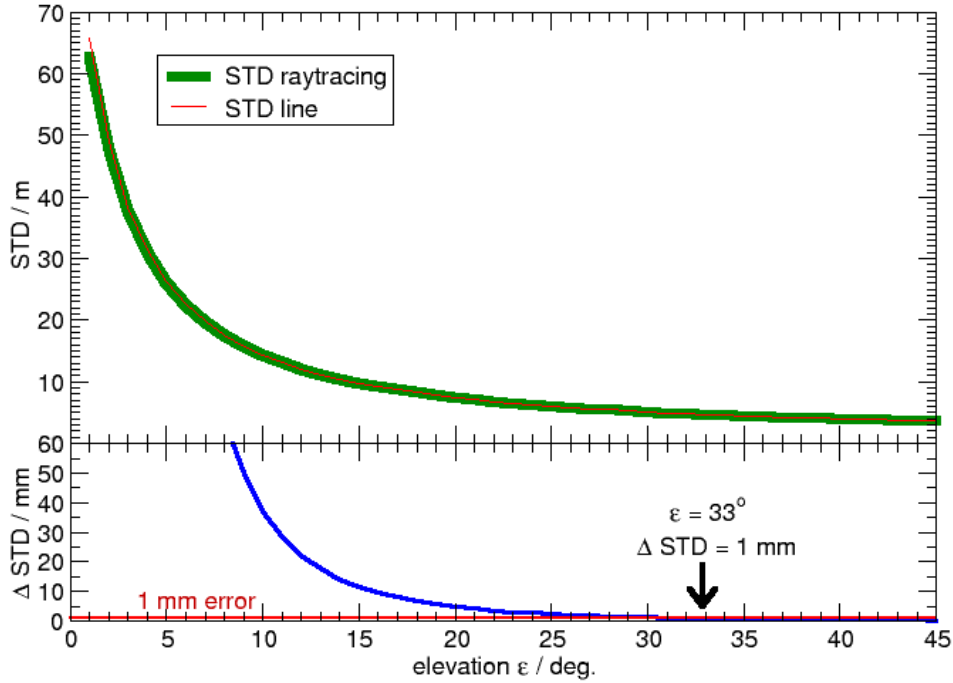
Figure 3.2: Raybendeing

The transformation ellipsoidal → cartesian coordinates is given by:

$$X = (N + h) \cdot \cos\beta \cdot \cos\lambda \tag{3.18}$$

$$Y = (N + h) \cdot \cos\beta \cdot \sin\lambda \tag{3.19}$$

$$Z = ((1 - e^2)N + h) \cdot \sin\beta \tag{3.20}$$

$$= \left(\frac{N}{1 + e'^2} + h\right) \cdot \sin\beta \tag{3.21}$$

$$= \left(\frac{b^2}{a^2}N + h\right) \sin\beta \tag{3.22}$$

These equations show no singularities and can be translated to computer code without any modifications.

The specific form of the transformation depends on the choice of axes and angles as given in ??. Here it is assumed that

$$0 \leq \lambda \leq 2\pi \tag{3.23}$$

$$-\frac{\pi}{2} \leq \beta \leq -\frac{\pi}{2} \tag{3.24}$$

The results are identical if the longitude is between $-\pi$ and $+\pi$, positive values indicating eastern directions and negative values indicating western directions. However, the transformation from cartesian to ellipsoidal coordinates always provides longitudes between 0 and 2π.

2) Transformations from cartesian to ellipsoidal coordinates   The transformation from cartesian to ellipsoidal coordinates is more complicated and several special conditions need to be regarded (routine `Cart2Ellips`):

$$\gamma \;=\; \arctan\left\{\frac{Z + e'^2\, b\sin^3\theta}{\rho - e^2\, a\cos^3\theta}\right\} \tag{3.25}$$

$$\varphi \;=\; 2\cdot\arctan\left(\frac{Y}{|X| + \rho}\right) \tag{3.26}$$

$$H \;=\; \frac{\rho}{\cos\beta} - N \tag{3.27}$$

with

$$\rho = \sqrt{X^2 + Y^2} \tag{3.28}$$

$$\theta = \arctan\frac{Z\,a}{\rho\,b}, \tag{3.29}$$

$$N = \frac{a}{\sqrt{1 - e^2\cdot\sin^2\beta}} \tag{3.30}$$

(β needs to be computed before N! Here, the latitude β is required, not φ!)

The ellipsoidal coordinates (λ, β, h) are:

$$\lambda \;=\; \begin{cases} 0° & \text{if } X = 0,\ Y = 0 \\ \varphi & \text{if } X \geq 0,\ Y \geq 0 \\ 360° + \varphi & \text{if } X \geq 0,\ Y < 0 \\ 180° - \varphi & \text{if } X < 0 \end{cases} \tag{3.31}$$

$$\beta \;=\; \begin{cases} +90° & \text{if } \rho = 0,\ Z > 0 \\ 0° & \text{if } \rho = 0,\ Z = 0 \\ -90° & \text{if } \rho = 0,\ Z < 0 \\ \gamma & \text{if } \rho \neq 0 \end{cases} \tag{3.32}$$

$$h \;=\; \begin{cases} -b & \text{if } \rho = 0,\ Z = 0 \\ |Z| - b & \text{if } \rho = 0,\ Z \neq 0 \\ H & \text{if } \rho \neq 0 \end{cases} \tag{3.33}$$

The longitude λ and the latitude β are within the range

$$0 \leq \ \lambda \ \leq 2\,\pi \tag{3.34}$$

$$-\frac{\pi}{2} \leq \ \beta \ \leq -\frac{\pi}{2} \tag{3.35}$$

the height h is givem in meters.

Transformations between the local horizon system and the "slant system"
The STD observations are given with respect to the local horizon system, i. e. the azimuth α and elevation ε of the GNSS satellite are given for a specific GNSS station. Azimuth

and elevation define a vector pointing from the GNSS station to the GNSS satellite and are independent from the bended signal path in the atmosphere.

The signal path can most easily be computed in a reference system which is aligned to the satellite-receiver-axis by solving Fermats's principle. Therefore, it is necessary to transform coordinates from the local horizon system to the "slant system". The "slant system" is defined as follows:

1. The origin of both reference systems is the GNSS station.

2. The x axis of the slant system points from the GNSS station to the GNSS satellite, i. e. defines the satellite-receiver-axis.

3. The z axis is restricted to a plane defined by the the center of Earth, the GNSS station and the GNSS satellite, i. e. for $\varepsilon = 0°$ the z axis is the vertical axis and for $\varepsilon = 90°$ the z axis is a horizontal axis.

4. The y axis is chosen in order to obtain a right handed Cartesian system.

It is assumed that azimuth and elevation are transformed to Cartesian coordinates defining a left handed local horizon system (x, y, z) (see XXX). The transformation from the local horizon system (x, y, z) to the slant system (x', y', z') is described by two basic rotations and one reflection at the y plane. The corresponding matrix R is given by

$$R = R_y(\varepsilon) \cdot R_z(\alpha) \cdot S_{xz} \tag{3.36}$$

with

$$R_y(\varepsilon) = \begin{pmatrix} \cos \varepsilon & 0 & \sin \varepsilon \\ 0 & 1 & 0 \\ -\sin \varepsilon & 0 & \cos \varepsilon \end{pmatrix}, \quad R_z(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S_{xz} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

The combined matrix

$$R(\varepsilon, \alpha) = \begin{pmatrix} \cos \varepsilon \cos \alpha & -\cos \varepsilon \sin \alpha & \sin \varepsilon \\ -\sin \alpha & -\cos \alpha & 0 \\ -\sin \varepsilon \cos \alpha & \sin \varepsilon \sin \alpha & \cos \varepsilon \end{pmatrix} \tag{3.37}$$

leads to a system of equations

$$\begin{aligned} x' &= x \cdot \cos \varepsilon \cos \alpha - y \cdot \cos \varepsilon \sin \alpha + z \cdot \sin \varepsilon \\ y' &= -x \cdot \sin \alpha - y \cdot \cos \alpha \\ z' &= -x \cdot \sin \varepsilon \cos \alpha + y \cdot \sin \varepsilon \sin \alpha + z \cdot \cos \varepsilon \end{aligned}$$

which is implemented in the subroutine `LocalHorz2Slant`, `STD_coord.f90`.

The inverse transformation from the slant system to the Cartesian local horizon system is given by

$$R^{-1} = R^{T} = S_{xz} \cdot R_z(\alpha) \cdot R_y(\varepsilon) \tag{3.38}$$

with

$$R^{-1}(\varepsilon, \alpha) = \begin{pmatrix} \cos\varepsilon\cos\alpha & -\sin\alpha & -\sin\varepsilon\cos\alpha \\ -\cos\varepsilon\sin\alpha & -\cos\alpha & \sin\varepsilon\sin\alpha \\ \sin\varepsilon & 0 & \cos\varepsilon \end{pmatrix} \tag{3.39}$$

The corresponding system of equations

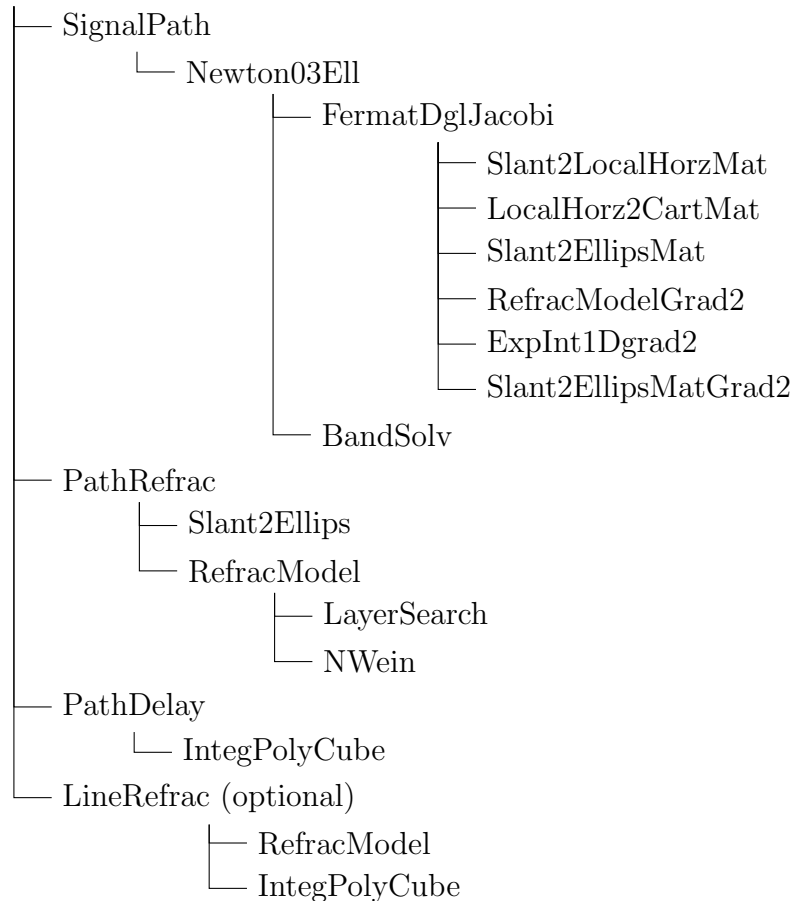$$
\begin{aligned}
x &= x' \cdot \cos\varepsilon\cos\alpha - y' \cdot \sin\alpha - z' \cdot \sin\varepsilon\cos\alpha \\
y &= -x' \cdot \cos\varepsilon\sin\alpha - y' \cdot \cos\alpha + z' \cdot \sin\varepsilon\sin\alpha \\
z &= x' \cdot \sin\varepsilon + z' \cdot \cos\varepsilon
\end{aligned}
$$

is implemented in the subroutine `Slant2LocalHorz`, `STD_coord.f90`.

> **Empty: GNSS Slant Total Delay operator: Implementation**

Subroutine structure of the STD operator

```
std_delay_1
        ├── SignalPath
        │       └── Newton03Ell
        │               ├── FermatDglJacobi
        │               │       ├── Slant2LocalHorzMat
        │               │       ├── LocalHorz2CartMat
        │               │       ├── Slant2EllipsMat
        │               │       ├── RefracModelGrad2
        │               │       ├── ExpInt1Dgrad2
        │               │       └── Slant2EllipsMatGrad2
        │               └── BandSolv
        ├── PathRefrac
        │       ├── Slant2Ellips
        │       └── RefracModel
        │               ├── LayerSearch
        │               └── NWein
        ├── PathDelay
        │       └── IntegPolyCube
        └── LineRefrac (optional)
                ├── RefracModel
                └── IntegPolyCube
```

COSMO Interface to the STD Operator

The main part of the COSMO interface to the STD observation oprator is the routine `organize_gnss_std` in file `src_obs_use_org.f90`.

The purpose of the COSMO interface is to

1. read and distribute the station data, e. g. coordinates, names,

2. read and distribute the STD observations and

3. collect and distribute the model fields (p, T, qv) required to compute the delays.

The station information needs to be read only once and is distributed to all nodes as the observations will be distributed in order to optimize the load balance and all nodes must be able to process observations from all stations. The STD observations are read as soon as new data are available. Currently, the STDs come in hourly batches which are read and distributed in one pass. Within each time step of the COSMO model it is checked if/how many STDs with the corresponding observation time are available. The required

module columns along the signal paths are requested and distributed between the nodes and the appropriate columns are transferred to the STD operator.

Part I

Appendix

Some Bibliography entries don't work. Somebody has to check this at some point...