Impact of atmospheric turbulence on geodetic very long baseline interferometry

T. Nilsson$^{1,2}$ and R. Haas$^1$

Received 29 April 2009; revised 1 October 2009; accepted 27 October 2009; published 12 March 2010.

We assess the impact of atmospheric turbulence on geodetic very long baseline interferometry (VLBI) through simulations of atmospheric delays. VLBI observations are simulated for the two best existing VLBI data sets: The continuous VLBI campaigns CONT05 and CONT08. We test different methods to determine the magnitude of the turbulence above each VLBI station, i.e., the refractive index structure constant $C_n^2$. The results from the analysis of the simulated data and the actually observed VLBI data are compared. We find that atmospheric turbulence today is the largest error source for geodetic VLBI. Accurate modeling of atmospheric turbulence is necessary to reach the highest accuracy with geodetic VLBI.


1. Introduction

Radio signals are delayed in the neutral atmosphere of the Earth. This delay is an important error source for space geodetic techniques, such as Global Navigation Satellite Systems (GNSS) and geodetic very long baseline interferometry (VLBI). We need to correct for the atmospheric delay in order to obtain the most accurate results with these techniques. Generally, the atmospheric delay is divided into two parts: one hydrostatic delay and one wet delay. The hydrostatic delay can be estimated using measurements of the atmospheric pressure at the surface. The wet delay can, however, not be estimated with high accuracy from surface measurements.

In VLBI mainly two methods exist to correct for the wet delay. The first one is to use an independent instrument, such as a water vapor radiometer (WVR), for estimating the atmospheric delay [Elgered et al., 1991; Kuehn et al., 1991; Emardson et al., 1999]. However, these estimates also contain errors, e.g., calibration errors and uncertainties in the conversion from measured brightness temperatures to atmospheric delays. Furthermore, a WVR does not give reliable results during rain.

The other method is to estimate the wet atmospheric delays along with the other parameters normally estimated in the VLBI data analysis, such as the station coordinates, clock errors, and Earth orientation parameters. This approach requires a model that describes the atmospheric delay as a function of observation direction and time. Typically, it is modeled as the delay in the zenith direction times a mapping function plus linear horizontal gradients [MacMillan, 1995]. This model assumes that the refractive index of air has only linear variations in the horizontal direction. The temporal variations in the model parameters are described either by piecewise linear functions (analysis based on classical least squares method [e.g., Ma et al., 1990]) or by random walk processes (analysis based on Kalman filtering [e.g., Herring et al., 1990]). However, such a model does not describe the small-scale fluctuations in the atmosphere caused by turbulence, since these are not linear.

The two methods give similar accuracy in the VLBI results [Emardson et al., 1999] despite problems to estimate wet atmospheric delays from WVR data observed during rain. There are two reasons for not achieving better results when estimating the atmospheric delay in the VLBI analysis compared to using an external calibration. The first one is increased sensitivity to noise since more parameters need to be estimated in the analysis. The other is that atmospheric turbulence is not represented in the atmospheric delay model used in the data analysis.

Here we present simulations designed to investigate the errors in the VLBI results due to atmospheric turbulence. We do this by simulating VLBI observations for the CONT05 and CONT08 campaigns (section 2), which are the two best existing VLBI data sets, using a method based on the theory of atmospheric turbulence (section 3). For the simulations we need the refractive index structure constant $C_n^2$, a parameter which in principle describes the magnitude of the turbulent variations. Several methods for estimating this parameter are presented in section 4. The simulation setup and the processing of the simulated and the actually observed VLBI data are described in section 5. The results and the conclusions are presented in sections 6 and 7, respectively.

2. CONT05 and CONT08 Campaigns

CONT05 and CONT08 are two VLBI campaigns, each consisting of 15 days of continuous observations, observed 12–27 September 2005 and 12–26 August 2008, respectively. Each campaign consisted of 15 one day ses-
Eleven VLBI stations participated in each campaign, of which nine were in both (see Figure 1).

Table 1 gives information about the station locations, the average number of scans observed at each station per day, the average number of stations involved in a scan, and the average number of unique radio sources observed by each station per day. As seen, there were more observations in CONT08 than in CONT05. This is because of shorter scans in CONT08 made possible by a higher data acquisition rate (512 Mbit s\(^{-1}\) compared to 256 Mbit s\(^{-1}\) in CONT05).

### 3. Theory

In this section we review the method for simulating atmospheric delays in a turbulent atmosphere [Nilsson et al., 2007; Nilsson and Haas, 2008]. Since turbulence affects both the wet and the hydrostatic parts of the total delays, we simulate the total delays. The slant total delay \( \ell_i \) of a radio signal observed at time \( t_i \) is given by

\[
\ell_i = \int_{S_i} [n(r(s)) - 1] \, ds,
\]

where \( S_i \) is the slant path of the signal and \( n(r) \) is the refractive index at position \( r \). The refractive index can be expressed as a function of the atmospheric pressure \( P \), temperature \( T \), and the partial pressure of water vapor \( e \) [Elgered, 1993]:

\[
n = 1 + 10^{-6} \left[ k_1 \frac{P}{T} + k_2 e + k_3 e^2 \right]. \tag{2}
\]

where \( k_1, k_2, \) and \( k_3 \) are constants.

When simulating atmospheric delays it is convenient to use the slant delay mapped to the zenith direction, the equivalent zenith delay \( \ell^*_i \):

\[
\ell^*_i = \frac{1}{m_i} \int_{S_i} [n(r(z)) - 1] \, ds = \int_{0}^{\infty} [n(r(z)) - 1] \, dz, \tag{3}
\]

where \( m_i = m(\epsilon_i) \) is the mapping function between elevation angle \( \epsilon_i \) and zenith. Typically, the mapping function is divided into two parts, one for the hydrostatic delay and one for the wet delay. Although turbulence of course causes fluctuations in both the wet and the hydrostatic delays, the fluctuations in the wet delay are typically a magnitude larger than those in the hydrostatic delay [Hill et al., 1988]. Hence,
we here use the wet mapping function of Niell [1996] for the fluctuating part of the total delay.

[11] In the presence of turbulence we can divide the refractive index into two parts, \( n = \bar{n} + \delta n \), where \( \bar{n} \) is the mean (nonturbulent) part and \( \delta n \) is the fluctuating (turbulent) part. Thus, \( \ell_{\delta} \) can also be written as a mean part \( \bar{\ell}_{\delta} \) plus a fluctuating part \( \delta \ell_{\delta} 
\]

\[
\ell_{\delta} = \int_0^\infty [\delta n(r(z), t)] \, dz + \int_0^\infty [\delta n(r(z))] \, dz = \bar{\ell}_{\delta} + \delta \ell_{\delta}.
\]

(4)

Most of the small-scale variations in \( \bar{\ell}_{\delta} \) will be caused by turbulence, i.e., the variations in \( \delta \ell_{\delta} \). Variations in the mean part \( \bar{\ell}_{\delta} \) are typically occurring slowly (timescales of several hours and more) and over large distances (>100 km). These are caused by, e.g., passages of weather fronts and evaporation/condensation of water vapor. We are mainly interested in the variations caused by turbulence; hence we will, for simplicity, assume that the mean part is constant.

The changes in the mean part should be modeled very well by the atmospheric model used in the VLBI analysis and thus will not cause any large errors in the VLBI results.

[12] We want to simulate a series of \( N \) equivalent zenith delays: \( \bar{\ell}_{\delta 0}, \bar{\ell}_{\delta 1}, \bar{\ell}_{\delta 2}, \ldots, \bar{\ell}_{\delta N-1} \). We choose one of these, e.g., \( \bar{\ell}_{\delta 0} \), as a reference and consider the difference between the other equivalent zenith delays and this one. We then calculate the covariance matrix \( C \), where the \((i, j)\)th component is [Nilsson et al., 2007]

\[
C_{ij} = \langle \delta \ell_{\delta i} - \delta \ell_{\delta 0} \rangle (\delta \ell_{\delta j} - \delta \ell_{\delta 0})
= \frac{1}{2} \int_0^\infty \int_0^\infty \left[ \langle \delta n(r(z), t_i) - \delta n(r(z), t_0) \rangle \right] \left[ \delta n(r(z), t_j) - \delta n(r(z), t_0) \rangle \right] \, dz \, dz,
\]

(5)

According to the theory of Kolmogorov turbulence, the structure function for the refractive index fluctuations is [Tatarskii, 1971; Treuhaft and Lanyi, 1987]

\[
\langle \delta n(r_i) - \delta n(r_j) \rangle^2 = C^2_n(z_i, z_j) \left[ \frac{\| r_i - r_j \|^{2/3}}{1 + \| r_i - r_j \|^{2/3}} \right].
\]

(6)

where \( C^2_n \) is called the refractive index structure constant, \( z_i \) is the vertical component of \( r_i \), and \( L \) is the saturation length scale (needed to avoid infinite variations at infinite distances). This equation, however, only gives the spatial variations in \( \delta n \). If we assume that the temporal variations are caused by the air moving with the wind (Taylor’s frozen flow hypothesis [Taylor, 1938]), then \( \delta n(r, t) = \delta n(r - v(t - t_0), t_0) \), where \( v \) is the wind velocity. This approximation is not valid for variations over long time periods, but should work reasonably well for short time variations (less than a few hours), which are of interest in this work. Hence, for spatial and temporal variations we have

\[
\langle [\delta n(r_i, t_i) - \delta n(r_j, t_j)]^2 \rangle = C^2_n(z_i, z_j) \left[ \frac{\| r_i - r_j - v(t_i - t_j) \|^{2/3}}{1 + \| r_i - r_j - v(t_i - t_j) \|^{2/3}} \right].
\]

(7)

Thus, if \( C^2_n \), \( L \), and \( v \) are known, it is possible to calculate the covariance matrix \( C \) given by equation (5). It is then possible to simulate \( \delta \ell = [\delta \ell_{\delta 1}, \ldots, \delta \ell_{\delta N-1}] \) simply by generating a series of \( N-1 \) random numbers with covariance matrix \( C \). In order to do this, we need to know the distribution of \( \delta \ell \). A simple case (which is a good approximation) is to assume a Gaussian distribution. In that case, we make a Cholesky factorization [Press et al., 1992] of \( C \)

\[
C = D D^T
\]

and obtain simulated values for \( \delta \ell \) by

\[
\delta \ell = \delta \ell_0 + Dw,
\]

(9)

where \( w \) is a vector of zero mean Gaussian random numbers with variance one. Then the simulated \( \ell' = [\bar{\ell}_{\delta 1}, \bar{\ell}_{\delta 2}, \ldots, \bar{\ell}_{\delta N-1}]^T \) will be obtained by (assuming \( \bar{\ell}_{\delta 0} = \bar{\ell} \) to be constant)

\[
\ell' = \bar{\ell} + \delta \ell = \bar{\ell} + \delta \ell_0 + Dw = \bar{\ell}_0 + Dw.
\]

(10)

A value for \( \bar{\ell}_{\delta 0} \) can be chosen freely, like a typical value for the zenith delay. Finally, in order to obtain the slant delays, the equivalent zenith delays are mapped to the slant direction by multiplying with the mapping function.

[13] For the simulations we need to know a number of parameters: the structure constant \( C^2_n \), the saturation length scale \( L \), the wind velocity \( v \), and \( \bar{\ell}_{\delta 0} \). As stated above, \( \bar{\ell}_{\delta 0} \) can be chosen as a typical value of the zenith delay. In this work we use 20 cm for all stations. The value used should not have any effect on the simulation results. The same is true for the value of \( L \), since the effect of \( L \) will only be seen at large scales (given that \( L \) is large). We use \( L = 3000 \) km, as used by Treuhaft and Lanyi [1987]. The wind velocity can be obtained, e.g., from numerical weather prediction models. We use a wind velocity constant with height and equal to the wind at the 850 hPa level obtained from European Centre for Medium Range Weather Forecasts (ECMWF) analysis data. The 850 hPa level (1400–1500 m altitude) corresponds approximately to the mean height of the atmospheric water vapor; thus, the wind at this level is a good choice for the wind to be used in the simulations. Methods for estimating \( C^2_n \) as well as some published values, are presented in section 4.4.

4. Methods for Estimating \( C^2_n \)

4.1. Radar Measurements

[14] One way to estimate a \( C^2_n \) profile is to use Doppler radar measurements. The power backscattered from a spe-
cific height is due to the irregularities caused by turbulence at that height. According to Otersten [1969], the radar reflectivity $\eta$ for clear air is related to $C_n^2$ through the expression

$$\eta = 0.38C_n^2\lambda^{-1/3},$$  \hspace{1cm} (11)$$

where $\lambda$ is the radar wavelength. According to Nastrom et al. [1986], the $C_n^2$ values determined from the radar should be accurate within a factor of 2.

[15] There exist a number of investigations where $C_n^2$ is measured with radars, although there are not many that present time series of $C_n^2$ in the lower troposphere from longer time periods (months to years). Chadwick and Moran [1980] presented 1 year of measurements of $C_n^2$ in Colorado. They found that the median value of $C_n^2$ at 805 m altitude varied from about $10^{-15}$ m$^{-2/3}$ in May–June to below $10^{-16}$ m$^{-2/3}$ in November. The observed values at this height were very variable, with occasional values larger than $10^{-13}$ m$^{-2/3}$. It was also shown that $C_n^2$ had a lognormal distribution and that log($C_n^2$) had a standard deviation of about 6 dB. Hence, the mean $C_n^2$ should be larger than the median ones by a factor of roughly 2.5.

[16] Nastrom et al. [1986] and Rao et al. [1997, 2001] used radar data to study the seasonal variations in $C_n^2$. They all found that $C_n^2$ generally is larger in summer than in winter. However, all these studies concentrated on higher altitudes (>4 km) where $C_n^2$ generally is much smaller than at lower altitudes.

[17] Other investigations have focused on shorter periods (a few days). Heo et al. [2003] showed estimated $C_n^2$ between 0 and 2.5 km above Viabon, France, for 2 days in June 1998. They found $C_n^2$ values ranging from $10^{-15}$ up to $10^{-12}$ m$^{-2/3}$, with most of the values being between $10^{-14}$ and $10^{-13}$ m$^{-2/3}$. In the various $C_n^2$ profiles presented by Cohn [1995], Hocking and Mu [1997], Luce et al. [1997], and Zink et al. [2004], $C_n^2$ is around $10^{-14}$ m$^{-2/3}$ at altitudes of 1–2 km.

4.2 Radiosonde Data

[18] It is possible to estimate $C_n^2$ using profiles of atmospheric pressure, temperature, humidity, and wind speed. Assuming that the small-scale fluctuations in the refractive index are created by turbulent mixing of large-scale refractive index variations, we can express $C_n^2$ as [VanZandt et al., 1978; d'Auria et al., 1993]

$$C_n^2 = a^2 F L_0^{4/3} M^2,$$  \hspace{1cm} (12)$$

where $a^2 \approx 2.8$, $L_0$ is the outer scale of atmospheric turbulence (proportional to the turbulence “mixing length”), $M$ is the potential refractive index gradient, and $F$ is called the intermittency factor expressing the fraction of the air that is turbulent. The potential refractive index gradient can be approximated by the gradient in the vertical direction since the vertical variations in the refractive index typically are much (more than 10 times) larger than those in the horizontal direction. This gradient includes only the part of the vertical gradient of $n$ caused by the vertical gradients in quantities conserved under adiabatic motion, i.e., the potential temperature $\theta$ and the specific humidity $q$:

$$\theta = T \left( \frac{P_0}{P} \right)^{0.286}$$  \hspace{1cm} (13)$$

$$q = e \left( \frac{1}{1.62T} \right).$$  \hspace{1cm} (14)$$

where $P_0 = 1000$ hPa. Since the specific humidity is a conserved quantity only for cloud-free conditions, this method cannot be used when clouds are present.

[19] Using equation (2) for the refractive index, we get

$$M = \frac{\partial \theta/\partial z}{\partial \theta/\partial z} \frac{\partial n/\partial z}{\partial q/\partial z}$$

$$= -10^{-6} \frac{P}{T} \left( k_i + 1.62 k_q \right) q \left( \frac{1}{\theta} \right) \frac{\partial q}{\partial z}$$

$$= -\left( 1.62 k_i + 1.62 k_q \right) \frac{\partial q}{\partial z}.$$  \hspace{1cm} (15)$$

The intervals over which the vertical gradients in $\theta$ and $q$ are calculated should be much larger than $L_0$, since these are the gradients in the mean quantities and should not contain any turbulence. Here we calculate $M$ over intervals of 500 m.

[20] The outer scale $L_0$ is generally unknown and can vary from less than 1 up to 100 m or more [d'Auria et al., 1993]. Typically, this problem is solved by assuming a statistical distribution for $L_0$ and replacing $L_0$ in equation (12) by $L_{0\text{eff}} = (L_0)^{4/3}$. Often $L_0$ is assumed to be uniformly distributed between $L_{0\text{min}}$ and $L_{0\text{max}}$. According to d'Auria et al. [1993], $L_{0\text{min}}$ is in the range of 0.1–5 m and $L_{0\text{max}}$ is in the range of 30–100 m. Here we assume $L_{0\text{min}} = 3$ m and $L_{0\text{max}} = 70$ m, i.e., approximately the midvalues of the given ranges. Using $L_{0\text{max}} = 100$ m instead of 70 m will cause $C_n^2$ to be larger by approximately a factor of 1.6.

[21] The intermittency factor $F$ is required since not all of the atmosphere is turbulent. This factor can vary between 0 and 1, typically $F = 0.1$ [VanZandt et al., 1978]. In this work, we have calculated $F$ using the Richardson number $R_i$:

$$R_i = \frac{g \partial \theta}{\partial z} \frac{\left| \partial q/\partial z \right|^2}{\partial q/\partial z}.$$  \hspace{1cm} (16)$$

One simple way to calculate $F$ would be to let $F = 1$ when $R_i$ is less than the critical Richardson number $R_i_c \approx 0.25$, and $F = 0$ otherwise. We have used a slightly more advanced model for calculating $F$ which gives a more smooth behavior of $F$ around $R_i = R_i_c$ [d'Auria et al., 1993; Vasseur, 1999]:

$$F = \int_{S_c}^{\infty} p(S) dS,$$  \hspace{1cm} (17)$$

where $S = \| \partial q/\partial z \|$, $S_c$ is the value of $S$ that makes $R_i = R_i_c$, and $p(S)$ is the probability density function for $S$ (assumed to be that of a Rice distribution; see d'Auria et al. [1993] for details).

[22] Figure 2 shows $C_n^2$ profiles obtained using high-resolution radiosonde data from Herstmonceux and St. Helena.
Table 3. Station Specific $C_n^2$ Values for CONT05 and CONT08 Obtained From Radiosonde Data and From the 24 h Variance of the Zenith Total Delay Observed by GPS

Table 2. Parameters of Exponential $C_n^2$ Profiles in Equation (18) for Different Seasons

4.3. Thermosondes

A thermosonde is a special type of radiosonde which, in addition to the measurements of pressure, temperature, humidity, etc., also measures with high accuracy (about $10^{-3}$ K) the variations in temperature over a small horizontal distance (typically around 1 m). These measurements are used to directly estimate the temperature structure constant $C_T^2$ [Azouit and Vernin, 2005]. Unfortunately, similar measurements of the humidity fluctuations are not made; thus, these will not give direct measurements of $C_n^2$. However,
profiles from thermosonde data calculated using equations (12) and (20).

Figure 3. $C_n$ profiles from thermosonde data calculated using equations (12) and (20).

since the temperature structure constant can be described by an expression similar to equation (12) as [Tatarskii, 1971]

$$C_n^2 = a^2 F L_n^{4/3} \left( \frac{P}{P_0} \right)^{0.286} \left( \frac{\partial T}{\partial z} \right)^2.$$  \hspace{1cm} (19)

we can relate $C_n^2$ to $C_n^2$ through

$$C_n^2 = M^2 \left( \frac{P}{P_0} \right)^{0.286} \left( \frac{\partial T}{\partial z} \right)^2 C_n^2.$$  \hspace{1cm} (20)

Hence, we no longer need to estimate $L_n$ and $F$ since these are embedded in $C_n^2$. Unfortunately, thermosonde launches are relatively rare.

[27] Figure 3 shows the mean $C_n^2$ profiles, calculated by equation (20) using $C_n^2_t$, pressure, temperature, and humidity profiles recorded by 16 thermosondes launched from Ash Mountain Helibase near Three Rivers, California, during the T-REX campaign from 20 March to 6 April 2006 [Jumper et al., 2007]. For comparison we show the mean $C_n^2$ profile calculated using equation (12), i.e., in the same way as done for the radiosonde data in section 4.2. We can see that the two methods agree relatively well.

[28] Zink et al. [2004] found that $C_n^2$ profiles estimated from radar and thermosondes agree within a factor of 2 after removing a bias. However, since they used the thermosonde profiles for calibration of the radar measurements, it is not possible to determine any possible bias between the two methods.

### 4.4. $C_n^2$ From WVR and GPS Measurements

[29] Measurements of the wet part of the atmospheric delay can also be used to estimate $C_n^2$. However, since the atmospheric delay is an integrated quantity it is not possible to infer the shape of the $C_n^2$ profile, so it must be assumed. Treuhaft and Lanyi [1987] made the simple assumption that $C_n^2$ was constant up to an effective tropospheric height $H$ and zero above. They obtained $C_n^2$ from measurements of the variance of the zenith wet delay $\ell$ over a time period $\tau$:

$$\sigma^2_{\ell}(\tau) = \frac{1}{\tau} \int_0^\tau (\tau - t) \left\{ (\ell(t + t_0) - \ell(t_0))^2 \right\} dt$$

$$= \frac{1}{2 \tau^2} \int_0^\tau (\tau - t) \int_0^H \int_0^H \left\{ \left[ (n(z,t_0) - n(z',t_0))^2 \right] + \left[ (n(z,t_0 + t) - n(z',t_0))^2 \right] - \left[ (n(z,t_0) - n(z',t_0))^2 \right] - \left[ (n(z,t_0 + t) - n(z',t_0))^2 \right] \right\} dz dz' dt$$

$$= \frac{1}{\tau} \int_0^\tau (\tau - t) \int_0^H C_2^2(z,z') \left( \left( z - z' \right)^2 + v^2 \tau^2 \right)^{1/3}$$

$$- \left| z - z' \right|^{2/3} dz dz' dt,$$  \hspace{1cm} (21)

where $t_0$ denotes the time of the beginning of the time period. Assuming $H = 1$ km, a mean wind speed of 8 m s$^{-1}$, and $\tau = 24$ h, they obtained a value of $C_n^2$ of $5.8 \times 10^{-12}$ m$^{-2/3}$ using zenith wet delay variances from WVR and radiosonde data. Later work, however, seems to indicate that using $H = 2$ km is more appropriate [Linfield and Wilcox, 1993; Keihm, 1995]. This gives $C_n^2 = 1.2 \times 10^{-12}$ m$^{-2/3}$.

[30] Using the same approach, we used the zenith total delay variances over 24 h obtained from GPS data to estimate $C_n^2$ for the CONT05 and CONT08 campaigns. The GPS data were analyzed with the GIPSY/OASIS-II Software [Webb and Zumberge, 1993], and the average of the 1 day variances was calculated for each station. Wind speeds were obtained from ECMWF (at the 850 hPa level). The derived values of $C_n^2$ are shown in Table 3. We estimated $C_n^2$ in two ways: By assuming $C_n^2$ is constant up to $H = 2$ km and by using a more realistic model:

$$C_n^2(z,z') = C_n^2 \exp \left[ - \frac{z + z'}{H} \right],$$  \hspace{1cm} (22)

where we have assumed $H = 2$ km (approximately the scale height of atmospheric water vapor). For both methods, we calculated one $C_n^2$ per station per day; then the average $C_n^2$ for each station was calculated.

[31] Nilsson et al. [2005] estimated $C_n^2$ from variations between slant wet delays of different directions, measured by a WVR. We have used the same method to estimate $C_n^2$ for the VLBI stations Onsala, Wettzell, and Kokee Park for the CONT05 period (Figure 4). Here we assumed that $C_n^2$ is constant up to $H = 2$ km and zero above. We obtain $C_n^2$ between $10^{-14}$ and $10^{-13}$ m$^{-2/3}$ for all three stations. As comparison the time series of $C_n^2$ estimated from GPS are also shown. As seen there is a good agreement for Onsala while the agreement is worse for Kokee and Wettzell. One reason for this could be that the sampling intervals of the Kokee and Wettzell radiometers are too large (45 s compared to around 10 s for Onsala) for the data to be really useful for estimation of $C_n^2$.

[32] A method for measuring atmospheric turbulence using data from a local network of GPS receivers was presented by Nilsson et al. [2009]. This method was applied to GPS data from the Yucca Mountain Network in Nevada. The results showed a strong seasonal variation in magni-
4.5. Summary of $C_n^2$ Measurements

[33] The review shows that $C_n^2$ is highly variable, both in time and as a function of location. Typically, we can expect the average $C_n^2$ for most locations to be in the range of $10^{-14}$–$10^{-13}$ m$^{-2/3}$ close to the ground and to be decreasing exponentially with height. We expect that $C_n^2$ is generally larger during warm and humid weather conditions.

5. Simulation Setup and Data Analysis

[34] Assuming that the station positions and velocities, the radio source positions, earth rotation, and orientation are known, and geophysical phenomena (such as solid earth tides, ocean loading, etc.) are modeled accurately, the observed minus the calculated group delay of a VLBI baseline observations can be described as

$$ o - c = \ell_1 - \ell_2 + \tau_1 - \tau_2 + w, $$

where $\ell_1$ and $\ell_2$ are the atmospheric delays for the two stations forming the baseline, $\tau_1$ and $\tau_2$ are the clock errors, and $w$ is the observation noise. Atmospheric delays are simulated using the procedure described in section 3. The observation noise is simulated as white noise with a standard deviation of 30 ps. A clock error is simulated as a random walk process plus an integrated random walk process plus an integrated random walk process. The standard deviation of the driving noise for these processes are $5 \times 10^{-7}$ ps s$^{-1/2}$, $1 \times 10^{-5}$ ps s$^{-3/2}$, and $5 \times 10^{-11}$ ps s$^{-5/2}$, respectively. These are typical values representative for a hydrogen maser (P. Jarlemark, SP Technical Research Institute of Sweden, personal communication, 2008).

[35] The simulated VLBI observations, as well as the actually observed VLBI data, were used in the VLBI analysis software SOLVE [Ma et al., 1990]. For each 24 h period we estimated the station positions (assumed constant over each 24 h period), the clock errors and the atmospheric zenith delays with a 1 h resolution, and the horizontal atmospheric gradients with a 6 h resolution.

6. Results

6.1. Same $C_n^2$ for All Stations

[36] We first consider the simple case of having the same $C_n^2$ profile at all stations. We made the simple assumption of a profile constant up to the height $H = 2$ km above ground and zero above. We tested three different values for $C_n^2$: $10^{-15}$, $10^{-14}$, and $10^{-13}$ m$^{-2/3}$. We expect the mean value of $C_n^2$ over, e.g., 1 day to be approximately $10^{-14}$–$10^{-13}$ m$^{-2/3}$ (see section 4).

[37] Figure 5 shows the baseline length repeatabilities obtained from the simulations. The repeatabilities obtained from analysis of the observed data are also shown for comparison. We note that for CONT08 the simulations with $C_n^2 = 10^{-13}$ m$^{-2/3}$ give repeatabilities which are on average similar to the observed ones; that is, the average $C_n^2$ for CONT08 was probably close to this value. For CONT05 $C_n^2 = 10^{-13}$ m$^{-2/3}$ gives on average too large repeatabilities while $C_n^2 = 10^{-14}$ m$^{-2/3}$ gives too small repeatabilities compared to the observed ones. Thus, the average $C_n^2$ for CONT05 was probably between $10^{-14}$ and $10^{-13}$ m$^{-2/3}$, i.e. smaller than for CONT08. This is reasonable since $C_n^2$ varies with season, and we expect that it has its maximum value in July–August in the northern hemisphere (in both campaigns, nine of 11 stations are in the northern hemisphere). Fur-
thermore, it is possible that Medicina and Zelenchukskaya (which participated only in CONT08) generally have larger \( C_n^2 \) than Algonquin Park and Gilmore Creek (which participated only in CONT05). Especially, we can expect a low \( C_n^2 \) at Gilmore Creek since it is located at a high latitude while the other stations are at midlatitudes (see Table 1).

The analysis of the simulations with \( C_n^2 = 10^{-15} \text{ m}^{-2/3} \) gives repeatabilities which are small compared to those obtained from the observed VLBI data. For such a small \( C_n^2 \) value, the errors in the VLBI results will mainly be from other error sources (i.e., clock errors and observation noise). Assuming that our simulations of the clock errors and observation noise are realistic (and that all other error sources can be ignored), the fact that the baseline repeatabilities obtained from the analysis of the observed data are much worse than those from the simulated data must be due to a value of \( C_n^2 \) that is too low. This shows that atmospheric turbulence is the largest error source for geodetic VLBI today.

6.2. \( C_n^2 \) From Radiosonde Data

We simulated atmospheric delays with \( C_n^2 \) as described by equation (18) and the model parameters given in Table 2. Unfortunately, the high-resolution radiosonde launch sites for which these parameters were obtained are not close to the VLBI sites that participated in CONT05 and CONT08. Hence, to obtain a \( C_n^2 \) profile for a specific VLBI station we used the profile from the radiosonde launch site closest in latitude. For CONT05 we used radiosonde profiles from the September–November period, and for CONT08 we used June–August.

The corresponding baseline length repeatabilities are shown in Figure 6. The repeatabilities from the simulations are generally larger than those from the observed data. There could be several reasons for this. One obvious problem is clearly that the \( C_n^2 \) profiles are not obtained close to the VLBI stations. Generally the difference in latitude is less than 10°; however, there could still be large differences in longitude (more than 150° in the worst case; see Tables 1 and 2). It could be that one or more of the radiosonde launch sites are located in a region with unusually high \( C_n^2 \). For example, it is known that \( C_n^2 \) can be large around and above steep mountains since a lot of turbulence is created when the wind is blowing toward the mountain. The radiosonde launch site at Gibraltar is such an example (we used the \( C_n^2 \) from Gibraltar for three VLBI stations). Another possibility is that the method for obtaining \( C_n^2 \) from radiosonde data is not accurate enough, since the uncertainty is factor of 2 (see section 4.2).

6.3. \( C_n^2 \) From the Variance of Zenith Delays

We also used \( C_n^2 \) profiles obtained from the 1 day variances of the zenith total delay estimated from GPS data acquired at the VLBI stations (Table 3). Simulations were made for both profile shapes (constant up to 2 km, and exponential). The baseline length repeatabilities from these simulations are presented in Figure 6.

As seen, the repeatabilities based on these simulations do agree rather well with those obtained from the observed VLBI data. However, for baselines in CONT05 containing Algonquin Park simulations clearly give worse baseline length repeatabilities compared with the observations. The method of using the 1 day variances of the zenith delay may not give realistic \( C_n^2 \) values for this station. One reason could be that there are significant nonturbulent variations. However, the method seems to work very well for
the other stations, at least to obtain $C_n^2$ that can be used in simulations.

We can also note that there are only small differences between using a constant $C_n^2$ up to 2 km and using a $C_n^2$ decreasing exponentially with height (equation (22)). This shows that the shape of the $C_n^2$ is of minor importance for the simulations. It is only important to know the integrated value.

Figure 7 shows the ratios between the baseline length repeatabilities obtained from the analysis of the simulated observations data and the real observations. It can be noted that the ratios are close to 1 for the simulations using the $C_n^2$ estimated from GPS. For CONT08 the mean values of the ratios are 0.94 and 1.0 for the simulations with $C_n^2$ constant with height and exponentially decreasing with height, respectively. For CONT05 the corresponding values are 1.3 for both cases. The baselines with Algonquin Park give the largest ratio values. The mean value of the ratios reduces to 1.15 when the baselines with Algonquin Park are excluded.

7. Conclusions and Outlook

The simulations presented indicate that the fluctuations in the atmospheric delays caused by turbulence presently are the largest error source for geodetic VLBI. Thus, in order to improve the accuracy of VLBI results the modeling of the atmosphere needs to be improved in the VLBI analysis and/or the observing strategy needs to be improved.

The simulation of VLBI observations can be a powerful tool to test any new VLBI observing and data analysis strategies. Atmospheric delay simulations using turbulence theory should be implemented into the scheduling software in order to improve and optimize observing schedules. This could easily be done using $C_n^2$ derived from GNSS data, reflecting, e.g., seasonal $C_n^2$ variations (section 6.3). The turbulence theory could also be implemented in the data analysis software. For example, the covariance matrix $C$ (equation (5)) can be used to weight the observations.

Another way to improve the VLBI results, even without introducing turbulence theory into the data analysis, is to increase the number of observations. An indication of such an improvement is that the baseline length repeatabilities for CONT08 are smaller than those for CONT05 in the simulations with constant $C_n^2$. The reason for the improved results is probably due to the fact that more observations with better sky coverage were made in CONT08 (see Table 1). We also note that the baseline repeatabilities for Algonquin Park for the simulations with $C_n^2$ from GPS are significantly higher than the repeatabilities for other baselines of similar length. This is because the $C_n^2$ values for Algonquin Park are high (see Table 3), but the number of scans is low (see Table 1). Other stations with high $C_n^2$ but with more observations, e.g., Westford in CONT05 and Tsukuba in CONT08, are not affected that much. For the observed data the baseline length repeatabilities are similar for CONT05 and CONT08. In this case the accuracy improvement due to more observations was compensated by degradation in accuracy because of higher $C_n^2$ values (see Table 3 and section 6.1).

The proposed future geodetic VLBI system, VLBI2010 [Niell et al., 2006; Behrend et al., 2008], will consist of radio telescopes with higher slew rates that will be
able to make more observations per unit time. The aim of the VLBI2010 system is to reach an accuracy of 1 mm in repeatability for global baselines. The methods presented in this work are currently used in simulations to evaluate whether and how this goal can be reached [see, e.g., Behrend et al., 2008; MacMillan, 2008; Wresnik et al., 2008].

[40] In order to do the simulations the structure constant C_n^2 needs to be known. As discussed in section 4, several methods exist to estimate this parameter. The problem is that the instruments used for many of these methods (e.g., radars and high-resolution radiosondes) are not available at very many places and in particular not at VLBI sites. To carry out realistic simulations, we need to know C_n^2 at the VLBI stations. The simulations indicate that the method of estimating C_n^2 from the variance of the zenith total delay has a sufficient accuracy and that the shape of the C_n^2 profile is of minor importance. Time series of the zenith total delay can be obtained from a number of different instruments, such as radiometers, GPS, or even VLBI data. For existing and planned VLBI stations, there are usually GPS stations nearby which can provide time series of zenith total delay for estimation of C_n^2.

[50] The method of simulating atmospheric delays presented in this work could, of course, also be used for simulations of other space geodetic techniques that use radio waves, such as GNSS and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS). In principle it should also be possible to use the method for simulating atmospheric delays of optical techniques, such as Satellite Laser Ranging (SLR). It should, however, be noted that C_n^2 for optical wavelengths is generally much lower than for radio waves.

[51] Acknowledgments. We are grateful to Gunnar Elgered for his valuable comments on the manuscript. The high-resolution radiosonde data were obtained from the British Atmospheric Data Center (BADC) (http://badc.necrc.ac.uk/data/rad-highres/). Thermosonde data were provided by NCAR/EOL under sponsorship of the National Science Foundation (http://data.col.eol.ucar.edu/). Wind data were obtained from ECMWF (http://www.ecmwf.int). The CONT05 and CONT08 campaigns were coordinated by the International VLBI Service for Geodesy and Astrometry (IVS) [Schlüter and Behrend, 2007]. This work was supported by the Swedish National Space Board.

References

Figure 7. Ratios between baseline length repeatabilities obtained from simulations and from observations. The results are from (top) CONT05 and (bottom) CONT08.


