RESUMO

A principal fonte de erro nas modernas técnicas geodésicas espaciais é a má modelação do atraso troposférico, denominação dada ao atraso sofrido pelas ondas rádio ao propagarem-se na camada electricamente neutra da atmosfera. De um modo geral, são consideradas duas componentes, a hidrostática (ou seca) e a húmida, cada uma das quais pode ser interpretada como o produto do atraso sofrido no zénite por uma função de mapeamento, que modela a dependência do ângulo de elevação do atraso troposférico.

Neste trabalho, é analisado o rigor obtido por quatro modelos de determinação da componente húmida do atraso troposférico. Usando dados de balões meteorológicos, a partir dos quais se geraram valores de referência, conclui-se que a exactidão destes modelos é, de um modo geral, insatisfatória. Assim, não é aconselhável a utilização isolada de modelos para correcção da componente húmida do atraso troposférico, na análise de dados de técnicas espaciais, se pretendermos obter precisão no posicionamento superior a alguns centímetros.

Zenith Wet Tropospheric Delay Determination Using Prediction Models: Accuracy Analysis

***V. B. Mendes*[1]**
***R. B. Langley*[2]**

ABSTRACT

The principal limiting error source in modern space geodesy techniques is the mismodeling of the delay experienced by radio waves in propagating through the electrically neutral atmosphere, usually referred to as the tropospheric delay. This propagation delay is generally split into hydrostatic (or dry) and wet components, each of which can be described as a product of the delay at the zenith and a mapping function, which models the elevation dependence of the propagation delay.

We assessed the performance of four prediction models developed to determine the wet zenith delay component of the tropospheric delay. Using ray tracing from radiosonde data as "truth", we concluded that the accuracy of these models is generally poor and that we can not rely on predictions from models to correct the effect of the tropospheric wet delay in the analysis of space geodetic data, if we want to achieve positioning accuracies at the few-centimetre level or better.

[1] *Faculty of Sciences of the University of Lisbon.*
[2] *Geodetic Research Laboratory, University of Brunswick.*

INTRODUCTION

The radio signals used by modern space geodetic techniques propagate through the earth's atmosphere — specifically through an ionized layer, the ionosphere, and a layer that is electrically neutral, composed primarily of the troposphere and stratosphere, referred to as the neutral atmosphere.

The ionosphere affects the propagation of the radio signals causing essentially a retardation of the modulation of the carrier wave, the **group delay**, and an advance of the carrier phase, the **phase advance**. These effects are equal in magnitude but opposite in sign. The ionospheric effect is proportional to the total electron content along the signal path and inversely proportional to the square of the frequency of the carrier. For dual frequency space techniques, this frequency-dependent property of the ionosphere (**dispersion**) allow us to combine linearly signals simultaneously transmitted on two frequencies and eliminate its major influence. The residual error can be neglected in most geodetic applications.

Unlike the ionized part of the atmosphere, the neutral atmosphere is essentially a nondispersive medium at frequencies below about 30 GHz (except for the anomalous

dispersion of the water vapor and oxygen spectral lines); i.e., the effects on phase and group delay are equivalent and the availability of more than one transmitted frequency is of no advantage in removing the tropospheric effect. Since the troposphere accounts for most of the neutral atmosphere mass and contains practically all the water vapor, the term **tropospheric delay** is often used to designate the global effect of the neutral atmosphere.

The neutral atmosphere affects the propagation of radio waves, causing a slowing down of the propagation velocity, resulting in a **propagation delay**, and, to a lesser extent, a **bending** of the ray path. These effects depend on the real-valued **refractive index**[1], n, along the signal ray path, more conveniently expressed by another quantity, the **refractivity**, N:

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

The refractivity of a parcel of air can be expressed, in general, as [Thayer, 1974]

$$N = K_1\left(\frac{P_d}{T}\right)Z_d^{-1} + \left[K_2\left(\frac{e}{T}\right) + K_3\left(\frac{e}{T^2}\right)\right]Z_w^{-1} \tag{2}$$

where P_d is the partial pressure of the dry gases in the parcel, e is the partial pressure of the water vapor, T is the absolute temperature, Z_d is the compressibility factor for dry air, Z_w is the compressibility factor for water vapor, and K_i are constants empirically determined. The two terms in the right-hand-side of equation (2) are the dry and wet components of the refractivity, respectively.

The compressibility factors are corrections to account for the departure of the air behavior from that of an ideal gas and depend on the partial pressure due to dry gases and temperature [Owens, 1967]. The most often used sets of refractivity constants are the ones provided by Smith and Weintraub [1953] and Thayer [1974] (see Table 1).

	Smith and Weintraub [1953]	Thayer [1974]
K_1	77.61±0.01	77.60±0.014
K_2	72±9	64.8±0.08
K_3	(3.75±0.03) · 10^5	(3.776±0.004) · 10^5

Table 1—Experimentally determined values for the refractivity constants (K_1 and K_2 are in K.mbar^{-1}, K_3 is in K^2. mbar^{-1}).

A simplified Smith and Weintraub two-term refractivity formula is widely used:

$$N = 77.6\left(\frac{P}{T}\right) + 3.73 \times 10^5\left(\frac{e}{T^2}\right). \tag{3}$$

[1] The imaginary part of the refractive index represents the absorption of a radio signal by atmospheric gases and liquid water.

This formula is estimated to have an accuracy of about 0.5%, for the range of temperatures normally encountered and for frequencies below 30 GHz. Thayer claims that his formula (eq. 2) represents an improvement of about one order of magnitude in the computation of N, yielding accuracies from about 0.02%, for dry air, to about 0.05%, for extremely moist air.

The dry component of the refractivity can be re-cast as a function of the total moist air density only, ρ, allowing its direct integration by applying the hydrostatic equation [Davis *et al.*, 1985]. As a consequence, the refractivity constant K_2 is also substituted with a new constant, K'_2, and the final expression for the refractivity can be given as a sum of a hydrostatic (as opposed to dry) and a wet component:

$$N = K_1 R_d \rho + \left[K'_2\left(\frac{e}{T}\right) + K_3\left(\frac{e}{T^2}\right)\right]Z_w^{-1} \tag{4}$$

where

$$K'_2 = \left(K_2 - K_1\frac{M_w}{M_d}\right) \tag{5}$$

and R_d is the specific gas constant for the dry air, M_w is the molar mass of water vapor, and M_d is the molar mass of dry air.

The tropospheric delay contribution d_{trop} to a radio signal propagating from a satellite to the earth's surface is given in first approximation by [Langley, 1992]:

$$d_{trop} = \int_{r_s}^{r_a} [n(r) - 1] \csc\theta(r)\, dr + d_{geo}, \tag{6}$$

where d_{geo} is the geometric delay that accounts for the difference between the refracted and rectilinear ray paths (ray bending), given by

$$d_{geo} = \left[\int_{r_s}^{r_a} \csc\theta(r)\, dr - \int_{r_s}^{r_a} \csc\varepsilon(r)\, dr\right], \tag{7}$$

and where r is the geocentric radius, θ is the refracted (apparent) satellite elevation angle, ε is the non-refracted (geometric or true) satellite elevation angle, r_s is the geocentric radius of the earth's surface, and r_a is the geocentric radius of the top of the neutral atmosphere.

This equation is valid for a spherically symmetric atmosphere, for which n varies simply as a function of the geocentric radius. The first integral in equation (6) represents the difference between the electromagnetic and geometric lengths of the refracted transmission path. The ray bending is explained by the Fermat's principle, which states that a ray follows the electromagnetically shorter path, and is significant at low elevation angles.

For a signal coming from the zenith direction, the geometric delay is zero; hence, equation (6) becomes, at the zenith:

$$d^z_{trop} = \int_{r_s}^{r_a} [n(r) - 1]\, dr = 10^{-6} \int_{r_s}^{r_a} N\, dr, \tag{8}$$

or, explicitly:

$$d^z_{trop} = 10^{-6} \int_{r_s}^{r_a} K_1 R_d \rho\, dr + \tag{9}$$

$$+10^{-6} \int_{r_s}^{r_a} \left\{ K'_2 \left(\frac{e}{T} \right) Z_w^{-1} + K_3 \left(\frac{e}{T^2} \right) Z_w^{-1} \right\} dr.$$

The delay just defined is the tropospheric *zenith delay*. It is convenient to express the tropospheric delay as two components: a hydrostatic component, associated with the dry molecular constituents of the atmosphere, and a wet component, associated with the water vapor in the atmosphere. Consequently, equation (9) can be written as:

$$d^z_{trop} = d^z_h + d^z_w \tag{10}$$

where d^z_h is the **hydrostatic zenith delay** and d^z_w is the **wet zenith delay.**

The hydrostatic zenith delay is very stable, as it is largely determined by surface pressure (see Fig. 1). It accounts for more than about 90% of the total delay and can be modeled using surface pressure measurements with an accuracy at the millimetre level [e.g. Janes et al., 1991], with the assumption that the atmosphere is in hydrostatic equilibrium. The Saastamoinen [1973] hydrostatic zenith delay model is frequently used (see Fig. 1), and is expressed, in metres, as:

$$d^z_h = 10^{-6} K_1 R_d \frac{P_s}{g_m}, \tag{11}$$

where P_s is the total surface pressure (mbar) and g_m is the magnitude of gravitational acceleration at the centroid of the atmospheric column (ms^{-2}), given by:

$$g_m = 9.784 \cdot (1 - 0.0026 \cos 2\varphi - 0.00028H) \tag{12}$$

where φ is the (geocentric) latitude of the station and H is the station orthometric height (km).

The wet component is a function of the atmospheric water vapor along the signal path. Unlike the hydrostatic delay, the wet delay is highly variable both in space and time and is a major residual error source in modern space geodesy techniques such as the Global Positioning System (GPS), very long baseline interferometry (VLBI), satellite altimetry (as featured, for example, on the SEASAT and TOPEX/Poseidon satellites), and Doppler Orbitography and Radiopositioning Integrated by Satellite (DORIS).

The strategies generally used to determine the wet zenith delay are (1) modeling using surface measurements of meteorological parameters; (2) «direct measurement» of the delay, using water vapor radiometers (WVRs); (3) estimation of the delay using the geometric strength of the radiometric data, along with other geodetic parameters, either as a constant or as a timevarying stochastic parameter.

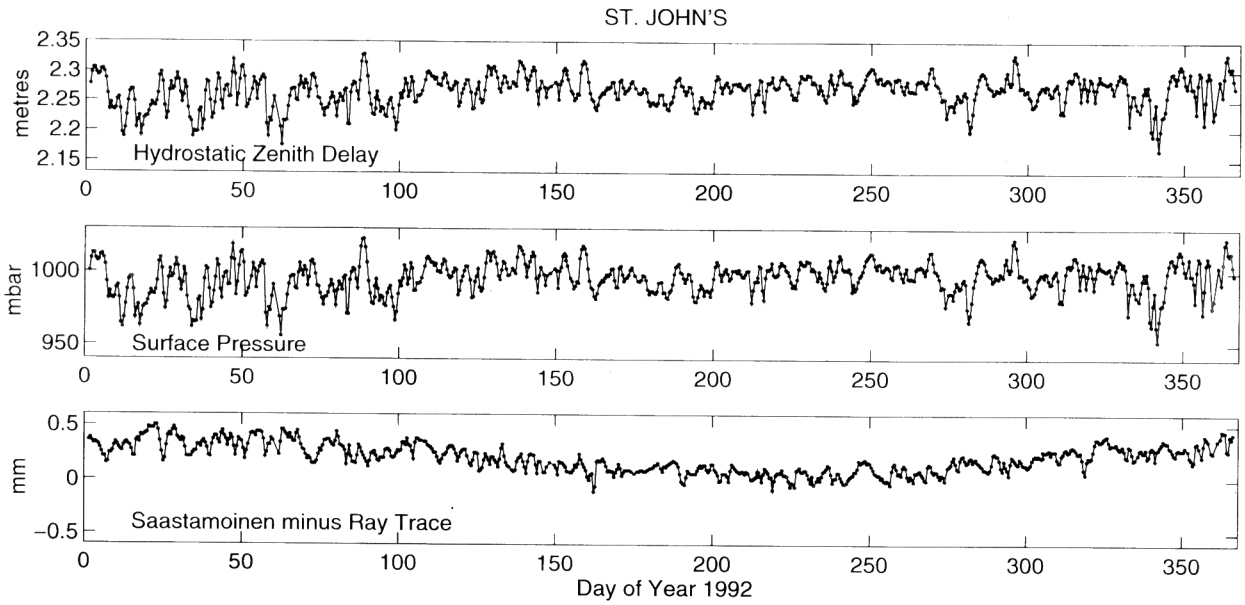

Figure 1 — Ray-traced hydrostatic zenith delay (top plot) and surface pressure (middle plot) for St. John's, showing the strong correlation between these quantities. The bottom plot gives an indication of how well the hydrostatic zenith delay can be determined using the Saastamoinen [1973] model (note the small annual error component).

Water vapor radiometers measure the brightness temperature (blackbody temperature) of the sky by using at least two radio frequencies close to the water vapor resonance spectral features (centered between 22 and 23 GHz) and provide a good indication of the temporal behaviour of the wet delay (for more details on the operation of WVRs see, for example, Elgered [1983] and Davis [1986]). They can calibrate the wet path delay, often with sub-centimeter accuracy, and are essential for studying the short-term variations (time-scales ranging from seconds to days) of the wet delay. Unfortunately, they are not readily available, they are expensive, and they can have a small bias, dependent on the site location and season.

The zenith delay can also be determined by ray tracing through radiosonde data (meteorological balloon measurements). The ray-tracing technique is helpful for investigating the long term variations of the wet delay, as meteorological balloons are regularly launched twice a day from many locations all over the world. These long term variations are mainly associated with the movement of air masses.

The zenith delay can be related to the delay that a transatmospheric radio signal would experience at other elevation angles through the use of **mapping functions**. In the early years of space geodesy, the zenith delay and the mapping function were not completely separated. With the evolution of the geodetic data analysis process, it became clear that it is more convenient to have a complete separation of these components, due to the generally poor modeling of the wet zenith delay from surface meteorological measurements. This separation would permit the zenith delay (or residual delay) to be estimated along with other geodetic or geophysical parameters from the space geodetic data. If the mapping functions are determined separately for the hydrostatic and the wet component, the tropospheric delay can be expressed as:

$$d_{trop} = d_h^z \cdot m_h(\varepsilon) + d_w^z \cdot m_w(\varepsilon) \tag{13}$$

where m_h, is the hydrostatic component mapping function, m_w is the wet component mapping function and ε, as above, is the non-refracted elevation angle at the ground station (some mapping functions use the refracted angle, θ). A comprehensive analysis of mapping functions used in the modeling the elevation dependence of the tropospheric delay has been given by Mendes and Langley [1994].

TROPOSPHERIC WET ZENITH DELAY PREDICTION MODELS

A very large number of zenith delay models have been developed in the last couple of decades. We have analysed three of the most recently developed. We also included the Saastamoinen [1973] wet delay model, which, despite its age, is used extensively as a prediction model in the analysis of space geodetic data. In the following discussion, T is the surface temperature (K), P is the total surface pressure (mbar), e is the surface water vapor pressure (mbar), U is the surface relative humidity (%), α is the temperature lapse rate (K/km), and d_w^z is the wet zenith delay (m).

The model by Saastamoinen is based on the assumption that the water vapor pressure decreases with height, similarly to the total pressure decrease, but more rapidly. For average conditions, Saastamoinen's wet zenith delay can be expressed as:

$$d_w^z = 0.002277\left(\frac{1255}{T} + 0.05\right)e. \tag{14}$$

Ifadis [1986] developed several zenith wet delay models, namely global, season dependent, and climatic dependent models. All the models are the result of regression analysis on an extensive set of radiosonde data and are based on the existence of a weak linear correlation of the zenith wet delay with the atmospheric pressure, water vapor pressure, and temperature. The global model analysed here is given by:

$$d_w^z = 0.00554 - 0.880\times10^{-4}\cdot(P-1000.0) + \\ + 0.272\times10^{-4}\cdot e + 2.771\cdot\left(\frac{e}{T}\right) \tag{15}$$

Askne and Nordius [1987] developed a two-parameter closed form model. These parameters, the temperature lapse rate α and the logarithmic scaling factor for the water vapor mixing ratio λ, model the typical variations of the temperature and humidity with height and should be chosen to fit the site and season. The model is given by the following expression:

$$d_w^z = 10^{-6}\frac{R_d e}{g_m(\lambda+1)}\left(K_2' + \frac{K_3}{T_m}\right) \tag{16}$$

where T_m is a mean temperature, given by:

$$T_m = T_s\left(1 - \frac{\alpha R_d}{g_m(\lambda+1)}\right) \tag{17}$$

In our analysis, we used the global value of λ (λ = 2.61) given in Smith [1966] and the standard value of the temperature lapse rate (6.5 K/km). However, it is important to realize that eq. (16) is especially sensitive to the variations in λ and α, which should be chosen to adequately fit the location and season.

Two types of models are proposed by Baby et al. [1988]. One of the types (theoretical) assumes that the relative humidity profile is constant and equal to its surface value up to a height H_m, where it reduces to zero. In addition, the temperature is assumed to decrease with increasing height at a constant rate and that the water vapor pressure is determined from:

$$e = e_s\frac{U}{100} \tag{18}$$

where the saturation pressure e_s is given by

$$e_s = \exp\left(A - \frac{B}{T} - C \ln T\right) \quad (19)$$

The coefficients A, B and C were determined through a least squares fit on laboratory measurements.

A semi-empirical model was also derived, given by the following expression:

$$d_w^z = 10^{-3} U v 10^{\gamma(T-273.15)}$$

where v and γ are empirical coefficients, which model the seasonal and climatic variations of the site. The authors claim that this model leads to a precision comparable to the theoretical model but with improved accuracy, and we have chosen this model for our analysis, using global values of v and γ ($v = 0.7284$ mm $(\%)^{-1}$ and $\gamma = 0.0236\ ^\circ C^{-1}$).

RAY TRACING

We used ray tracing to provide benchmark values for our analysis. To compute the zenith wet delay to be used as «truth», we used radiosonde data from eight stations (see Table 2) representing different climatic regions. The radiosonde data pertains to the year 1992, each station having typically two balloon launches per day, at 11 h and 23 h UT (Guam, San Juan, Nashville, Oakland, Denver) and 0 h and 12 h UT (for the other stations). The balloon meteorological data consists of height profiles of pressure, temperature, and relative humidity or dew-point (Landvetter). Radiosonde profiles with obvious errors (e.g. no surface data recorded) were discarded. Table 3 shows the statistics for the hydrostatic and wet zenith delays obtained for each station.

STATION	φ(°N)	λ (°W)	H(m)	NP
Guam	13.55	215.17	111	736
San Juan	18.43	66.10	3	675
Nashville	36.12	86.68	180	745
Oakland	37.73	122.20	6	740
Denver	39.75	108.53	1611	753
St. John's	47.62	52.75	140	713
Landvetter	57.67	347.70	155	714
Alert	82.50	62.33	66	720

Table 2 — Approximate locations of the radiosonde sites and the corresponding number of profiles (NP) used. H is the height of the station above the geoid.

	HYD (m)		WET (m)	
STATION	mean	rms	mean	rms
Guam	2.281	0.009	0.274	0.062
San Juan	2.316	0.005	0.264	0.045
Nashville	2.270	0.012	0.151	0.081
Oakland	2.313	0.010	0.115	0.035
Denver	1.908	0.012	0.073	0.040
St. John's	2.268	0.026	0.092	0.056
Landvetter	2.262	0.027	0.166	0.009
Alert	2.282	0.022	0.032	0.023

Table 3 — Mean and root-mean-square about the mean of the ray traces for the hydrostatic (HYD) and wet components of the zenith delay for the different radiosonde stations.

RESULTS AND CONCLUSIONS

The most significant fact resulting from this analysis is the large and similar scatter about the mean shown by all models (see Table 4 and Fig. 2 and 3). The overall best performance with respect to ray-tracing values is presented by both the Ifadis and Sabbath models, which exhibit a very strong similarity. The Askne and Nordius model has a tendency to overestimate the zenith delay in the presence of high water vapor content whilst the Baby et al. model seems to underestimate it. The performance of these models can however be optimized for a particular station. As mentioned before, the models are very sensitive to the choice of the parameters that drive the model. Notice, for example, a very small bias for Nashville, shown by the Askne and Nordius model, which may reflect «tuned» parameters for this station.

We conclude that the determination of the wet zenith delay can not be obtained with an accuracy better than about 2 cm, using any of these models. However, due to the large scatter shown by all models, this accuracy is likely unrealistic and errors of several centimeters in the determination of the zenith wet delay should be expected. The Ifadis and Saastamoinen models seem, nevertheless, to have an advantage over the Askne and Nordius and Baby et al. models and these models are recommended.

	Askne and Nordius		Baby et al.		Ifadis		Saastamoinen	
Station	mean	rms	mean	rms	mean	rms	mean	rms
Guam	38	45	-26	47	-5	46	1	46
San Juan	32	33	-28	34	-9	33	-3	33
Nashville	1	31	-24	34	-16	31	-17	31
Oakland	32	28	8	26	14	27	15	27
Denver	2	19	-8	17	10	18	-7	18
St. John's	-4	33	-12	34	-12	34	-15	33
Landvetter	15	22	4	23	6	22	3	22
Alert	-8	10	-5	9	-6	8	-11	8

Table 4 — Mean and root-mean-square (rms) scatter about the mean of the differences between the zenith wet delay predictions of the models and the ray-trace results, in millimetres. The rms values represent, in most of the cases, more than about 20% of the mean wet zenith delay.

ACKNOWLEDGMENTS

We would like to express our appreciation to Arthur Niell for providing the ray-trace software (developed by James Davis, Thomas Herring, and Arthur Niell).

Radiosonde data was provided by Environment Canada (Atmospheric Environment Service), U.S. National Center for Atmospheric Research, and the Swedish Meteorological and Hydrological Institute.

The support of *PRAXIS XXI/Junta Nacional de Investigação Científica e Tecnológica* is gratefully acknowledged.

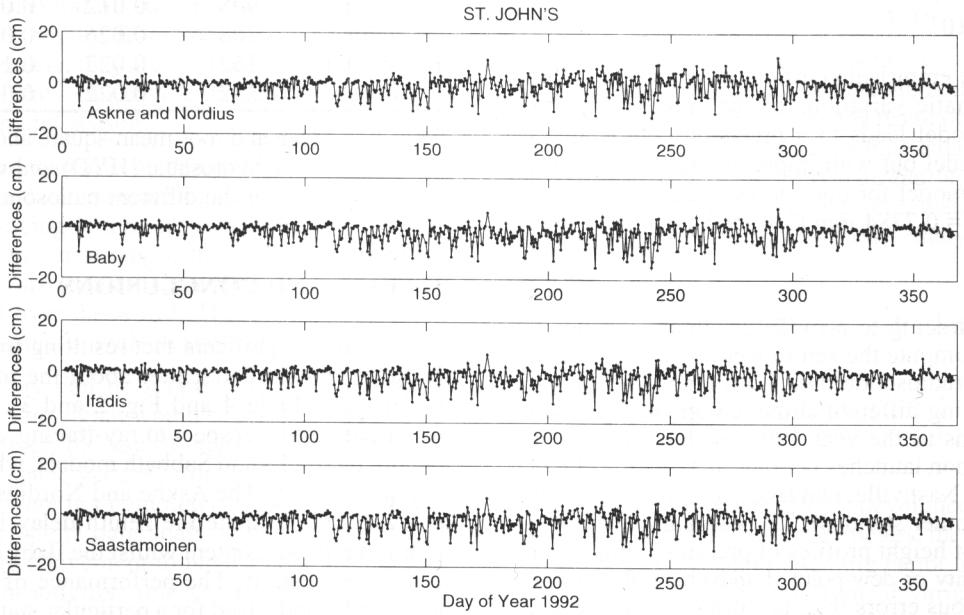

Figure 2 — Differences between the wet zenith delay computed using the prediction models and the ray-traced results, for St. John's (low to moderate wet zenith delay).

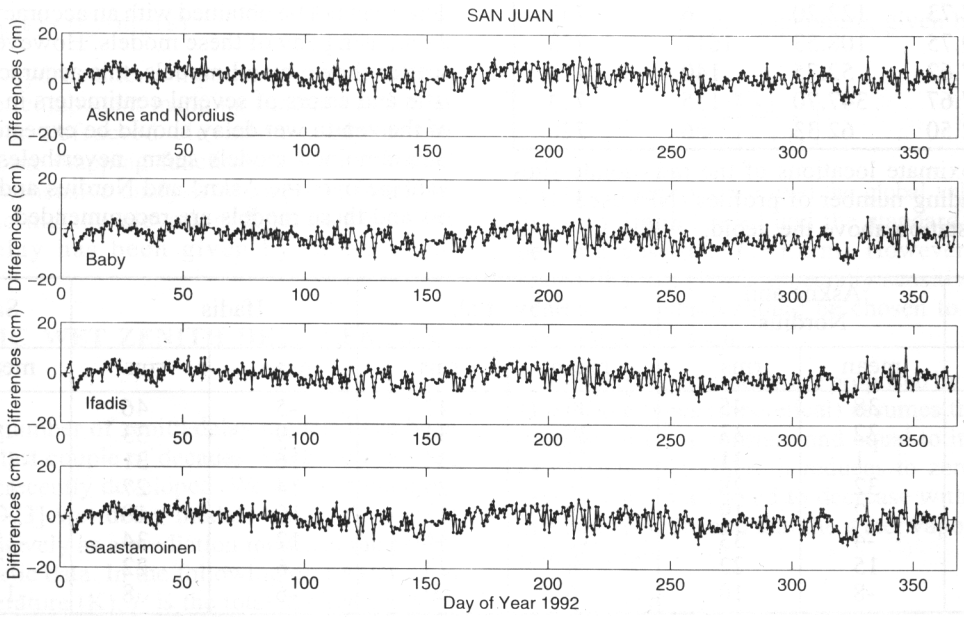

Figure 3 — Differences between the wet zenith delay computed using the prediction models and the ray-traced results, for San Juan (high wet zenith delay).

REFERENCES

Askne, J. and H. Nordius (1987). «Estimation of tropospheric delay for microwaves from surface weather data.» *Radio Science*, Vol. 22, No. 3, pp. 379-386.

Baby, H.B., P. Golé, and J. Lavergnat (1988). «A model for the tropospheric excess path length of radio waves from surface meteorological measurements.» *Radio Science*, Vol. 23, No. 6, pp. 1023-1038.

Davis, J.L., T.A. Herring, I.I. Shapiro, A.E.E. Rogers, and G. Elgered (1985). «Geodesy by radio interferometry: Effects of atmospheric modeling errors on estimates of baseline length.» *Radio Science*, Vol. 20, No. 6, pp. 1593-1607.

Davis, J.L. (1986). «Atmospheric propagation effects on radio interferometry.» Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA; U.S. AFGL Technical Report 860243, U.S. Air Force Geophysical Laboratory, Hascom AFB, Mass.

Elgered, G. (1983). *Water Vapor Radiometry With Applications to Radio Interferometry and Meteorology*, Ph.D. Thesis, Chalmers University of Technology, Goteborg, Sweden.

Ifadis, I.I. (1986). «The atmospheric delay of radio waves: Modeling the elevation dependence on a global scale.» Technical Report 38L, Chalmers University of Technology, Goteborg, Sweden.

Janes, H.W., R.B. Langley, and S.P. Newby (1991). «Analysis of tropospheric delay prediction models: Comparisons with ray-tracing and implications for GPS relative positioning.» *Bulletin Géodésique*, Vol. 65, pp. 151-161.

Langley, R.B. (1992). «The effect of the ionosphere and troposphere on satellite positioning systems.» Paper presented at the Symposium on Refraction of Transatmospheric Signals in Geodesy, The Hauge, The Netherlands, May 19-22.

Mendes, V.B. and R.B. Langley (1994). «A comprehensive analysis of mapping functions used in modeling tropospheric propagation delay in space geodetic data.» KIS94, International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, Banff, Canada, 30 August - 2 September, 1994, pp. 87-98.

Owens, J.S. (1967). «Optical refractive index of air: Dependence on pressure, temperature, and composition.» *Applied Optics*, Vol. 6, No. 1, pp. 51-59.

Saastamoinen, J. (1973). «Contributions to the theory of atmospheric refraction.» In three parts. *Bulletin Géodésique*, No. 105, pp. 279-298; No. 106, pp. 383-397; No. 107, pp. 1334.

Smith, W. (1966). «Note on the relationship between total precipitable water and surface dew point.» *Journal of Applied Meteorology*, Vol. 5, pp. 726-727.

Smith, E.K. and S. Weintraub (1953). «The constants in the equation of atmospheric refractive index at radio frequencies.» *Proceedings of the Institute of Radio Engineers*, Vol. 41, No. 8, pp. 1035-1037.

Thayer, G.D. (1974). «An improved equation for the radio refractive index of air.» *Radio Science*, Vol. 9, No. 10, pp. 803-807.