In: de Munck, J.C., Spoelstra, T.A.Th. (Eds.): Refraction of Transatmospheric Signals in Geodesy. Netherlands Geodetic Commission, Publications on Geodesy 36, 1992

MONITORING TOTAL IONOSPHERIC ELECTRON CONTENT AND IONOSPHERIC IRREGULARITIES WITH GPS

LAMBERT WANNINGER
Institut für Erdmessung (IfE), Universität Hannover
Nienburger Str. 6, 3000 Hannover, F.R.G.

ABSTRACT. Dual-frequency code and carrier phase GPS measurements offer an excellent opportunity to monitor total ionospheric electron content (TEC) and ionospheric irregularities. This paper will discuss observation biases and errors. It will be shown that differential equipment group delays of the satellite and receiver hardware are the most significant biases. Their accurate determination depends on the ionospheric conditions. The ability to monitor irregularities is not affected by these biases.

1. IONOSPHERIC OBSERVABLES

The ionosphere is defined as that part of the atmosphere where sufficient ionization can exist to affect the propagation of radio waves. The most important effects on GPS signals are the retardation, or group delay, on the modulation carried on the radio wave, and the advance of the carrier phase. The degree of both effects depends on the physical properties of the ionosphere and the frequency of the penetrating electromagnetic wave – the higher its frequency the less the effect. The group delay of a signal propagated through the ionosphere is directly proportional, at least to first order, to total ionospheric electron content (TEC) along the signal path. First order effects of the phase advance correspond to those of the group delay except for their sign (e.g. Klobuchar, 1985). Higher order effects do not exceed 0.5 % of TEC at GPS frequencies, but they are usually much smaller. Therefore, TEC can be estimated from measurements of the two L-band GPS signals.

The ionospheric electron content is integrated along the signal path up to the GPS satellite altitude of 20,000 km. TEC from GPS includes the plasmaspheric electron content above an altitude of some 2000 km. Its contribution to the total electron content amounts to about 10–50 % mainly depending on the time of day and the world region (Soicher, 1977; Davies, 1990). Other TEC measuring systems (Faraday rotation measurements to geostationary satellites, or dual-frequency Doppler measurements to NNSS Transit satellites) or ionospheric models (Bent model, IRI) do not take the plasmaspheric electron content into account.

TEC is calculated from GPS code observations by

$$TEC_P = S * (P_2 - P_1) \quad [el/m^2] \tag{1}$$

$$S = \frac{1}{40.3} * \frac{f_1^2 * f_2^2}{f_1^2 - f_2^2} = 9.52 * 10^{16} \quad [el/m^3] \tag{2}$$

with the GPS frequencies $f_1 = 1575.42$ MHz and $f_2 = 1227.60$ MHz, S being a factor to convert differences of dual-frequency GPS P-code measurements P_1, P_2 $[m]$ to TEC $[el/m^2]$.

TEC can also be determined from dual frequency phase measurements except for the unknown numbers of whole cycles, the carrier phase ambiguities N, by

$$TEC_\phi = -S * \left((\phi_2 + N_2) * \lambda_2 - (\phi_1 + N_1) * \lambda_1\right) \tag{3}$$

with subscripts identifying the signals L_1 and L_2, ϕ [cycles] being the measured carrier phases, λ [m] the carrier wavelength, and S is defined in equation (2). Phase measurements are only suitable for determination of TEC if the unknown carrier phase ambiguities can be estimated.

Another way to use the ionospheric information of dual frequency phase measurements is to eliminate the ambiguities by forming time differences. If no cycle slip occurs between two measurement epochs t_1 and t_2, the time difference is free of any unknowns:

$$\Delta TEC_\phi(\Delta t_{12}) = -S * \Big((\phi_2(t_2) - \phi_2(t_1)) * \lambda_2 - (\phi_1(t_2) - \phi_1(t_1)) * \lambda_1\Big) \tag{4}$$

The resolution of GPS carrier phase measurements is so high that this rate of TEC observable gives an excellent means to observe ionospheric irregularities. ΔTEC_ϕ depends very much on the GPS satellite orbits for two reasons. On the one hand the elevation angle of the signal incidence direction changes with up to 0.5 deg/min. The smaller the elevation angle the longer is the signal path through the ionosphere and the larger is the measured TEC. Thus, the primary effect which is observed is the change of the elevation angle. However, periodic effects caused by irregularities can be detected quite easily, as long as their period is considerably smaller than the length of an orbit arc as seen from one station (2–7 hours, 1000–2000 km in ionospheric altitude). Moreover, the point of interception of a GPS signal with the area of ionospheric irregularities (in an altitude of e.g. 400 km) moves due to the satellite orbit with a velocity of 50 to 450 m/s. Taking into account that a typical GPS receiver has a maximum output rate of 1 Hz, an irregularity has to have an extension of at least some hundred meters to be observable. Smaller structures can be detected but not observed in detail. In this paper rate of TEC over a time difference of 1 min is discussed because it can be obtained as a by-product of geodetic GPS measurements.

2. BIASES AND ERRORS

In order to estimate the electron content of the ionosphere differences of dual-frequency measurements are formed. All those biases and errors which are the same for both frequencies cancel out: clock biases of the satellites and of the receiver, clock dithering due to Selective Availability (SA), and tropospheric refraction. The remaining biases and errors are differential equipment group delays, multipath, phase center offsets, phase center variations and imaging effects.

2.1 Differential equipment group delays

The P-codes transmitted from GPS satellites at the two frequencies show a synchronization bias due to different hardware paths inside the transmitter. Similar effects are also known for the receiver hardware.

The requirements on the GPS satellites demand that the satellite differential equipment group delays shall not exceed 15 ns ($43*10^{16} el/m^2$) and random variations shall not exceed 3 ns ($9*10^{16} el/m^2$, 2σ) (Rockwell, 1984). Actual differential delays were determined in prelaunch factory tests. These values do not exceed 3 ns ($9*10^{16} el/m^2$) for Block I satellites (Dahlke et al., 1988; Coco et al., 1991). Prelaunch testing values are also included in the satellite broadcast message (Rockwell, 1984). They have been broadcast for all satellites since the fall of 1990. Maximum values are about $12*10^{16} el/m^2$. But these values do not agree with those in the literature, nor with values estimated as combined satellite and receiver differential delays at Institut für Erdmessung (IfE) from GPS P-code measurements.

Satellite and receiver differential delays can only be separated from each other if additional ionospheric information is introduced: Coster and Gaposchkin (1989) separated these delays by using an ionospheric model, and thus they introduced errors due to inaccuracies of the model.

From GPS observations alone an estimation of combined satellite and receiver differential delays can be performed as shown by Lanyi and Roth (1988) and Dahlke et al. (1988). They used TEC_P measurements from a mid-latitude nighttime observation period to estimate the coefficients of a two-dimensional low-order polynominal model of the vertical TEC and the combined satellite and receiver differential delays in a least-squares process.

The same approach can be used for TEC_ϕ data (Wild et al., 1989). Here, one unknown is estimated for each satellite consisting of the carrier phase ambiguities plus satellite and receiver differential equipment phase delays. One additional unknown has to be introduced for each cycle slip. But accurate results are only achievable if (almost) all cycle slips have been removed beforehand. The advantage of using TEC_ϕ-data lies in the ability to use observations of hybrid receivers, which do not measure the P-code on L_2 but obtain the phase on L_2 by squaring the signal.

Four important assumptions are made in this estimation method: TEC is constant in a rotating reference frame, TEC and vertical TEC (VTEC) are related by an obliquity factor which is a function of the elevation angle only, satellite and receiver differential delays are constant over several hours and the electron distribution of the ionosphere shows small horizontal gradients in order to fit to a simple model.

Therefore, best results can be achieved in a mid-latitude nighttime observation period. Results are poorest in equatorial daytime observation periods with their large TEC gradients. The main problem is the separation of the receiver differential delay from average VTEC. Both unknowns are highly correlated due to the limited geometry of GPS-TEC-measurements. Measurements at low elevation angles reduce this correlation but introduce the problems of increased multipath (can be effectively reduced by carrier smoothed code) and of much poorer results in mapping TEC to VTEC. The contribution of the satellite differential delays to the combined delays is estimated with much higher accuracy. Here, a dense coverage of the sky with satellite paths is of importance. Mid-latitude and high-latitude sites experience a poorer coverage than equatorial sites because of the Block II satellite orbital inclination of just 55 deg.

If satellite and receiver delays or combined delays are known, GPS-TEC_P-observations can immediately be corrected. But this requires that the delays are fairly constant between two calibrations. Differential equipment group delays are known to be very sensitive to temperature changes. The receiver must be calibrated at the temperature of its later use. Geodetic GPS-receivers are usually not built to keep a constant differential equipment delay. Periodic effects or drifts do not affect positioning accuracy, as long as they are identical for all channels. Experiences with Ashtech P-code receivers (beta version) show that day-to-day variations of up to 3 ns ($9*10^{16} el/m^2$) occur. Variations of the TI 4100 are much smaller if the receiver is temperature controlled.

Coco et al. (1991) reported that no significant variations of differential delays of Block I satellites could be found over a time span of five weeks, nor over a time span of 2 years. No experiences with variations of differential delays of Block II satellites have been reported in the literature.

If the variations of the differential delays are too big or as long as they are not calibrated, combined delays must be determined from existing GPS observations together with a simple ionospheric model. This approach succeeds in areas and time periods of very small horizontal gradients (mid-latitude, nighttime) with an accuracy of about $2*10^{16} el/m^2$. But in areas and time periods of large gradients (e.g. equatorial anomaly, daytime) the accuracy degrades considerably.

2.2 Multipath

Multipath errors occur if the received signal is composed of the direct line of sight signal and one or more constituents which have propagated along paths of different length. Different propagation paths result from either reflection at the satellite or in the surroundings of the receiving antenna. Multipath effects are not identical for signals of different frequencies. Thus, they do not cancel out forming an ionospheric observables. The effect of multipath on code measurements is 2 orders of magnitude larger than its effect on carrier phase measurements (Bishop et al., 1985).

Satellite multipath can hardly be proven nor estimated from actual GPS measurements (Young et al., 1985). Its effect is expected to be considerably smaller than maximum receiver multipath.

Experiments at IfE showed that the effect of receiver multipath on ionospheric observables can reach up to 10 ns (code) or 0.1 ns (phase) in a highly reflective environment. It appears in cyclic variations with periods from 1 min to more than 1 h due to changing geometry between satellite, reflector and antenna. Especially the short periodic multipath effects influence the rate of TEC observable with biases of up to 0.1 ns/min. Multipath effects can be minimized by selection of an appropriate antenna (reduced low-angle response), careful attention to antenna siting and the use of absorbing material in the vicinity of the antenna. Effects on TEC_P observations can substantially be reduced by combining code and carrier observations.

2.3 Phase center offsets and imaging

The physical center of a GPS antenna does not generally coincide with the point at which the signal is received. The point to which the radio measurements are referred, the phase center, is the apparent electrical center of the antenna. The phase centers of the two GPS frequencies are independent of each other and will coincide only by chance. Moreover, the apparent phase center position will be a function of the signal

TABLE 1.
Biases and errors of ionospheric GPS observables

Observable	TEC_P	TEC_ϕ	$\Delta TEC_\phi(\Delta t = 1$ min$)$
Units	$10^{16}el/m^2$	$10^{16}el/m^2$	$10^{16}el/(m^2 * min)$
Range	1 - 500	1 - 500	0 - 7
Differential equipment group delays			
- of satellites	0 - 12 (- 43)		
- of receiver	0 - 20 (- ??)		
- combined determination from observations			
(including carrier phase ambiguities of TEC_ϕ)	2 - 15	2 - 15	
Multipath	0 - 30	0 - 0.3	0 - 0.3
Phase center offsets		0 - 0.5	0.00
Random Observation Errors (1σ)	1.5 - 15	0.02 - 0.05	0.03 - 0.07

incidence direction due to a non-spherical phase pattern of the antenna. The antenna phase patterns of the two frequencies are independent of each other.

Phase center locations for copies of a single model of antenna tend to be consistent. They can be determined in laboratory experiments and then be used to correct the ionospheric observable. Schupler and Clark (1991) determined phase center offsets and phase center variations for 5 standard GPS antennas, according to which L_1/L_2 offsets can reach 0 – 5 cm ($0.0 - 0.5*10^{16}el/m^2$). The full bias propagates to the ionospheric observable when the L_1/L_2 offset direction coincide with the signal incidence direction. Rate of TEC will almost not be affected because the L_1/L_2 offset changes slowly with time due to the slow changes in satellite geometry.

Antenna imaging is caused by conducting material in the vicinity of the antenna. An antenna 'image' in the conductor interferes with the phase pattern of the antenna (Tranquilla, 1986). It causes environmentally-induced phase center variations. This effect diminishes with increasing distance between antenna and conducting body. Antenna imaging has to be avoided by careful attention to antenna siting.

2.4 Random observation errors

Observation errors are due to the limitations of the receiver's electronics and are of a random character. They can effectively be reduced by averaging. Depending on the quality of the receiver a single P-code measurement can be performed with an observational error (1σ) of 0.1 – 1 m and a carrier phase measurement with 1 – 3 mm. These errors propagate to the ionospheric observables with $1.5 - 15*10^{16}el/m^2$ and $0.02 - 0.05*10^{16}el/m^2$ respectively. The rate of TEC observable has then a random observation error of $0.03 - 0.07*10^{16}el/m^2$ per time unit. The accuracy of codeless (cross-correlating) receivers is degraded at low elevation angles.

2.5 Summary

Biases and errors of ionospheric GPS observables are summarized in Table 1. Differential equipment group delays are of systematic character. If they are not corrected, they can easily lead to misinterpretation of TEC data, especially when TEC is small. The accuracy of a combined determination from observations depends very much on the ionospheric conditions. Multipath effects and random observation errors of code measurements can be reduced to a neglectable level by smoothing with the carrier phases. Multipath is the only error source of importance of the rate of TEC observable. It must be avoided by careful attention to antenna site selection.

3. CONCLUSIONS

The accuracy of GPS-TEC-observations is primarily limited by differential equipment group delays originated in the satellite and receiver hardware. An estimation of the combination of these biases can be performed from GPS observations. But the accuracy of this determination depends on the present conditions of the ionosphere. Therefore, it would be very advantageous to use predetermined calibration values of these biases. This requires receivers with a very much stable differential delay. It requires as well the observation of the satellite differential delays and their publication. GPS rate of TEC observations offer en excellent means to monitor ionospheric irregularities. Here, the effect of biases and errors are usually very small.

ACKNOWLEDGEMENT. This work was supported by the Deutsche Forschungsgemeinschaft under grant SE 313/12.

REFERENCES

Bishop, G.J., Klobuchar, J.A., Doherty, P.H. (1985). Multipath effects on the determination of absolute ionospheric time delay from GPS signals. *Radio Science*, **20**, 388-396

Coco, D.S. (1991). GPS – Satellites of opportunity for ionospheric monitoring. *GPS World*, **2**, No.9, 47-50

Coco, D.S., Coker, C., Dahlke, S.R., Clynch, J.R. (1991). Variability of GPS satellite differential group delay biases. *IEEE Transactions on Aero.Elec.Sys.*, **27**, 931-938

Coster, A.J., Gaposchkin, E.M. (1989). Use of GPS pseudo-range and phase data for measurement of ionospheric and tropospheric refraction. *Proc.ION-GPS '89*, Colorado Springs

Dahlke, S.R., Coco, D.S., Coker, C.E. (1988). Effect of GPS system biases on differential group delay measurements. ARL-TP-88-17, The University of Texas, Austin

Davies, K. (1990). *Ionospheric Radio*. Peter Peregrinus Ltd., London

Georgiadou, Y., Kleusberg, A. (1988). On carrier signal multipath effects in relative GPS positioning. *Manuscripta Geodaetica*, **13**, 172-179

Klobuchar, J.A. (1985). Ionospheric time delay effects on earth-space propagation. In *Handbook of Geophysics and the Space Environment*, ed. A.S. Jursa, U.S. Air Force

Lanyi, G.E., Roth, T. (1988). A comparison of mapped and measured total electron content using GPS and beacon satellite observations. *Radio Science*, **23**, 483-492

Rockwell (1984). NAVSTAR GPS space segment / navigation user interfaces ICD-GPS-200, Rockwell Int. Space System Group

Soicher, H. (1977). Plasmaspheric signal time-delay effects in satellite navigation systems. In *AGARD CP 209*

Schupler, B.R., Clark, T.A. (1991). How different antennas affect the GPS observable. *GPS World*, **2**, No.10, 32-36

Tranquilla, J.M. (1986). Multipath and imaging problems in GPS receiver antennas. *Proc.4th Int.Geod.Symp.Sat.Pos.*, Austin

Young, L.E., Neilan, R.E., Bletzacker, F.R. (1985). GPS satellite multipath: an experimental investigation. *Proc.1stInt.Symp.Pos.GPS*, Rockville

Wild, U., Beutler, G., Gurtner, W., Rothacher, M. (1989). Estimating the ionosphere using one or more dual frequency GPS receivers. *Proc.5thInt.Geod.Sym.Sat.Pos.*, Las Cruses