Regional Stochastic Models for NOAA-Based Residual Tropospheric Delays

Hassan E. Ibrahim and Ahmed El-Rabbany

(*Ryerson University*)
(Email: rabbany@ryerson.ca)

Real-time and near real-time precise GPS positioning requires shorter GPS solution convergence time. Residual tropospheric delay, which exists as a result of the limitations of existing tropospheric correction models, is a limiting factor for quick GPS solution convergence. This paper proposes a new approach to tropospheric delay modelling, which overcomes the limitations of existing models. In this approach, the bulk of the tropospheric delay is accounted for using the NOAA-generated tropospheric correction model, while the residual tropospheric delay component is accounted for stochastically. First, the NOAA tropospheric correction model is used to generate daily time series of zenith total tropospheric delays (ZTDs) at ten IGS reference stations spanning North America for many days in 2006. The NOAA ZTDs are then compared with the new highly-accurate IGS tropospheric delay product to obtain daily residual time series at 5 minute intervals. Finally, the auto-covariance functions of the daily residual tropospheric delay series are estimated at each of the ten reference stations and then used to find the best empirical covariance function in the least squares sense. Of the three potential covariance functions examined, it is shown that the exponential cosine function gives the best fit most of the time, while the second-order Gauss-Markov model gives the worst fit. The first-order Gauss-Markov fits are close to those of the exponential cosine. Additionally, the model coefficients seem to be season independent, but change with geographical location.

KEY WORDS

1. Tropospheric delay.   2. NOAA tropospheric corrections.   3. Stochastic models.

1. **INTRODUCTION.** Tropospheric path delay is typically expressed as a product of unknown total zenith path delay (ZPD), which is modelled as a random walk process noise, and a known mapping function relating slant path delay to ZPD (Kouba and Héroux, 2001). Alternatively, the tropospheric path delay can be modelled as a function of zenith hydrostatic and wet delays, with two different mapping functions, plus tropospheric gradients (Gao et al., 2004). In such a case, the zenith wet delay and two gradient coefficients are to be estimated. Unfortunately, unlike the zenith hydrostatic delay, the zenith wet delay is highly correlated with the total zenith tropospheric delay (Gutman et al., 2003), which in turn is known to be highly correlated with the height component of the station coordinates

Figure 1. IGS station distribution in North America. (From http://igscb.jpl.nasa.gov/network/maps/all_nam.html).

(Kouba, 2003). The existence of such high cross-correlation as well as the temporal correlation of the tropospheric path delay slows down the convergence of the GPS solution. More recently, a number of regional and local monitoring networks have been established to generate tropospheric corrections. Among them is the NOAA tropospheric correction model, which incorporates GPS observations into numerical weather prediction (NWP) models (Gutman et al., 2003). The model is distributed as a 2-D grid file, which is updated hourly and contains the zenith tropospheric delay (current and forecast) over the U.S. and surrounding regions (including a large portion of Canada). Unfortunately, however, although the model improved the positioning solution compared with the standard tropospheric models (e.g., Saastamoinen and Hopfield models), a residual tropospheric delay component remains un-modelled. Such a residual delay component is found to be temporally correlated, which may slow down the conversion of the GPS solution.

To overcome the limitations of the current models, a new approach to tropospheric modelling is proposed in this paper. In this approach, the bulk of the tropospheric delay is accounted for using the NOAA-generated tropospheric correction model, while the residual tropospheric delay component is accounted for stochastically. To do this, daily residual tropospheric delay data series at ten International GNSS Service (IGS) reference stations are obtained by comparing the NOAA tropospheric model and the new final tropospheric product of the IGS. The new IGS final tropospheric delay product, which is generated at a 5-minute interval, is described as highly accurate and highly reliable (Byun et al., 2005). The ten IGS reference stations included in this research are ALGO, AMC2, FLIN, HLFX, HOLB, JPLM, MOD1, NLIB, PRDS and USNO (Figure 1). The stations are selected in different geographical locations in order to represent the various climate conditions in our proposed model. Also the seasonal changes have been considered in the model by using data over the year 2006.

Estimates of a set of auto-covariance functions of the un-modelled daily tropospheric delay residuals are obtained at each of the ten reference stations, which are used to develop the empirical covariance functions, in the least-squares sense. Three

empirical covariance functions are examined as potential covariance functions, namely first- and second-order Gauss-Markov processes and the exponential cosine function. Fitting results show that the exponential cosine function gives the best fit most of the time, while the second-order Gauss-Markov model gives the worst fit. The first-order Gauss-Markov fits are close to those of the exponential cosine. Additionally, the model coefficients seem to be season independent, but change with geographical location.

2. NOAA TROPOSPHERIC DELAY MODEL. The US National Oceanic and Atmospheric Administration (NOAA) tropospheric corrections model has been developed by the NOAA Forecast Systems Lab (FSL) (Gutman et al. 2003). The model is based on numerical weather prediction (NWP) models, where surface- and space-based meteorological measurements and others are combined into the model (Ahn et al., 2006). The NOAA model estimates the zenith hydro-static (dry) tropospheric delay as follows:

$$ZHD = (2 \cdot 768 \pm 0 \cdot 0024) \times 10^3 \frac{P_{sfc}}{f(\varphi, H)} \tag{1}$$

where, ZHD is the zenith hydrostatic tropospheric delay, H is the station orthometric height obtained from a digital elevation model (DEM), φ is station latitude and P_{sfc} is the surface pressure given by:

$$P_{sfc} = [Alt^{0 \cdot 193} - (1 \cdot 313 \times 10^{-5})H]^{5 \cdot 255} \tag{2}$$

where, Alt is the altimeter setting in mbar, which is calculated by using the surface pressure and the station height H from a geopotential surface that approximates the mean sea level, and

$$f(\phi, H) = (1 - 0 \cdot 00266 \, cos2\varphi - 0 \cdot 00028H). \tag{3}$$

The NOAA model also estimates the zenith tropospheric wet delay from the integrated precipitated water (IPW) retrieved from the weather model as follows:

$$ZWD = 10^{-6} \left(\frac{k_3}{T_m} + k_2' \right) \times R_v \times IPW \tag{4}$$

$$T_m = 70 \cdot 2 + 0 \cdot 72 T_{sfc} \tag{5}$$

$$k_2' = k_2 - \left(\frac{R_d}{R_v} \right) k_1 \tag{6}$$

$$\left. \begin{array}{l} k_1 = 77 \cdot 6 \pm 0 \cdot 05 \, K/hPa \\ k_2 = 70 \cdot 4 \pm 2 \cdot 2 \, K/hPa \\ k_3 = (3 \cdot 739 \pm 0 \cdot 02)10^5 \, K^2/hPa \\ R_d = 287 \cdot 06 \, JKg^{-1} \, K^{-1} \\ R_v = 461 \cdot 525 \, JKg^{-1} \, K^{-1} \end{array} \right\} \tag{7}$$

where, ZWD is the zenith wet tropospheric delay, T_{sfc} is the surface temperature, T_m is the weighted mean temperature of the atmosphere, R_d and R_v are the dry and wet air constants, respectively, and k_1, k_2, k_3 are refraction constants (see Ahn et al., 2006 for more details).

NOAA has developed a software package that generates text files of hourly analyzed and 3-hour predicted zenith dry, wet and total tropospheric delays. The NOAA tropospheric delay software consists of a number of codes that interact to compute the dry, wet and total tropospheric delays as final products. More information about the software can be found at (ftp://aftp.fsl.noaa.gov/gpsmet/zwdgrids/tools).

3. EMPIRICAL COVARIANCE FUNCTIONS. As indicated earlier, daily data series of residual tropospheric delays obtained by comparing the new IGS tropospheric product with the NOAA ZTD were used to generate estimates of auto-covariance function (ACF) at different times and geographical locations. Assuming that the data represents a stationary random process, the unbiased ACF can be defined as:

$$R_{xx}(\tau) = \frac{1}{N - |\tau|} \sum_{t=0}^{N-|\tau|-1} x(t)x(t+\tau) \tag{8}$$

where, τ is the lag or shift, $x(i)$, $i = 0, 1, 2, \ldots, $ N-1 is the data sequence with the mean removed, and N is the series number of points. As we deal with individual stations, we used the normalized ACF, which is obtained by dividing the ACF at lag τ and the ACF at zero lag.

Since the ZTD residuals time series are of limited length, the reliability of the estimated auto-covariances decreases as the lag increases. As such, we only used the first 10% of the auto-covariance estimates to develop the empirical auto-covariance function (El-Rabbany, 1994).

Three functions have been tested as potential empirical covariance functions for the stochastic model. The best model was selected as the one with the least root-mean-square (RMS) error, in the least-squares sense. The first model is the first-order Gauss-Markov given by:

$$f(\tau) = exp\,(-\beta|\tau|) \tag{9}$$

where, β is the unknown model coefficient (reciprocal of the correlation time), which can be determined using, for example, least-squares estimation, and τ is the time lag. The second model is the second-order Gauss-Markov given by:

$$f(\tau) = exp\,(-\beta|\tau|)(1+\beta|\tau|) \tag{10}$$

where, β is the model coefficient (2·1461945/the correlation time), which again can be determined using the least-squares estimation, and τ is the time lag. The third model is the exponential cosine function given by:

$$f(\tau) = exp\,(-\beta|\tau|)\,cos(\alpha|\tau|) \tag{11}$$

where, β and α are positive quantities representing the unknown parameters of the model.
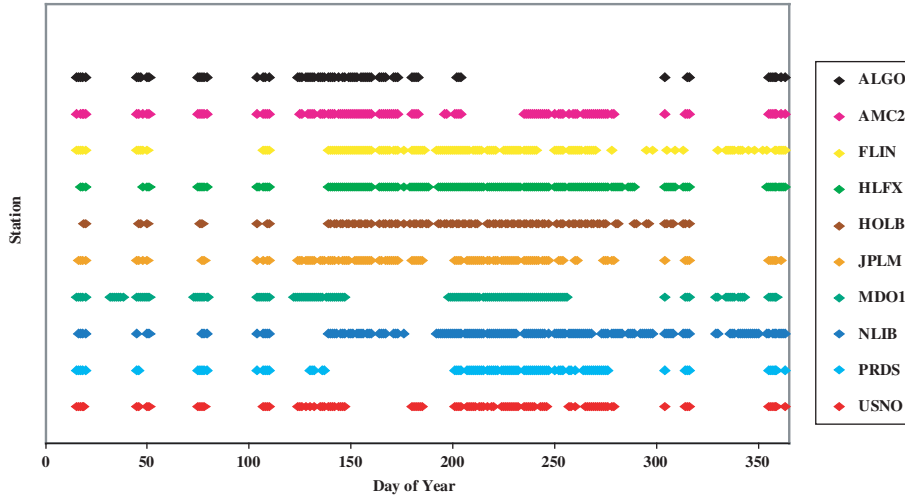
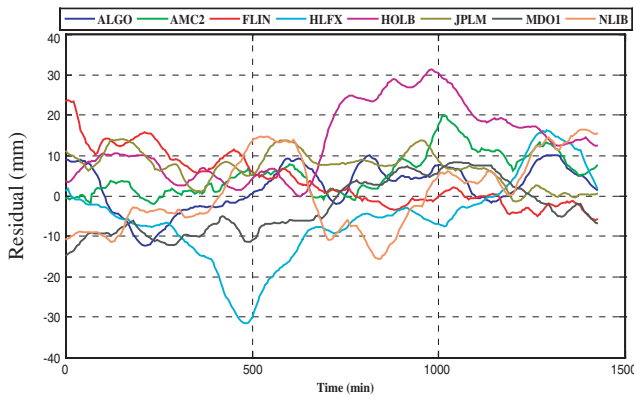Figure 2. Distribution of days processed for each station in the year 2006.

Figure 3. NOAA-based residual tropospheric delays at 8 stations for day 20 of the year 2006.

4. RESULTS AND DISCUSSION. A total of 1350 daily *ZTD* data series for the ten reference stations were generated from the new IGS and NOAA tropospheric correction products. The *ZTD* data series are distributed over many days of the year 2006, representing the various seasons (Figure 2). While the new IGS tropospheric delay product is generated at a 5-minute interval, the NOAA software package produces one record every hour for each of the *ZHD*, *ZWD* and *ZTD*. To generate *ZTD* residuals at a 5-minute interval, we interpolated all the NOAA-based *ZTD* series to 5 minutes using the Lagrange interpolation method. Figures 3 through 6 show examples of the NOAA-based residual tropospheric delay series (i.e., NOAA-based *ZTD* minus IGS-produced *ZTD*) for four days representing different seasons of the year.

Figures 7 and 8 show the variations of the NOAA-based residual tropospheric delays for stations AMC2 and USNO over the year 2006. As can be seen in Figures
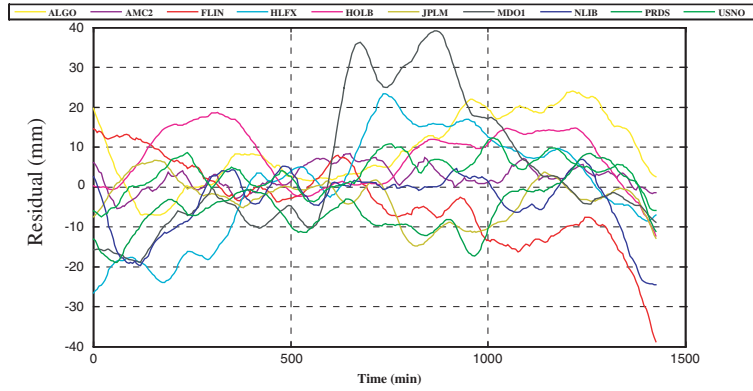
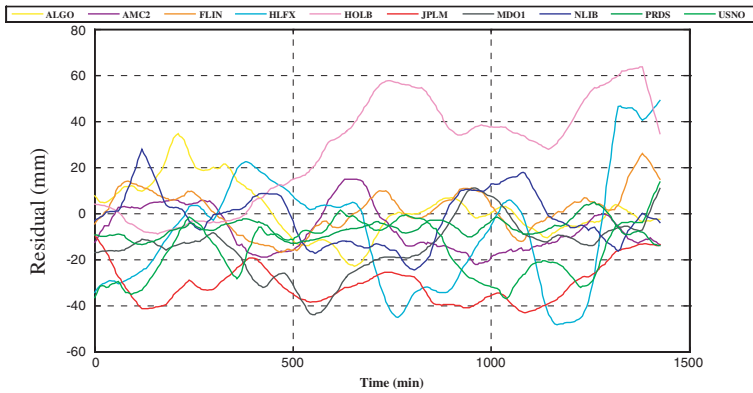Figure 4. NOAA-based residual tropospheric delays at 10 stations for day 110 of the year 2006.

Figure 5. NOAA-based residual tropospheric delays at 10 stations for day 202 of the year 2006.
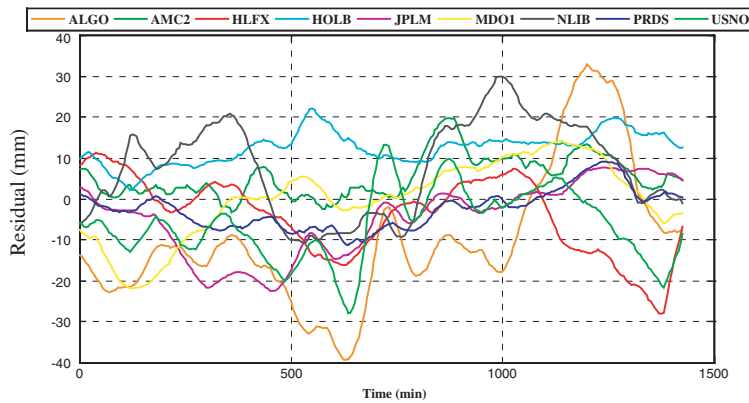
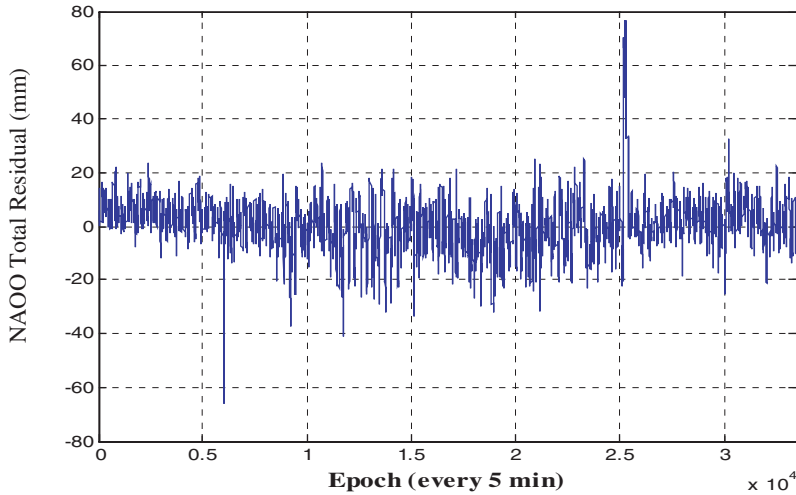Figure 6. NOAA-based residual tropospheric delays at 9 stations for day 304 of the year 2006.

Figure 7. NOAA-based residual tropospheric delays over 117 days of the year 2006 for AMC2 station.
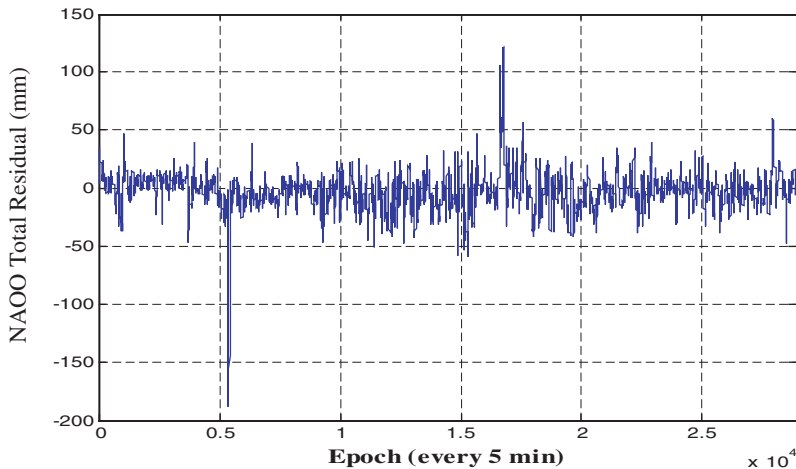

Figure 8. NOAA-based residual tropospheric delays over 101 days of the year 2006 for USNO station.

7 and 8, apart from a few spikes, the NOAA-based residual tropospheric delays can be considered to have no significant variations over the year. This indicates that the NOAA model accounts for the seasonal variations quite well. Consequently, a single stochastic model can be developed for a particular location, regardless of the time of the year. Investigation of the reason for the spikes in the NOAA-based residual tropospheric delays revealed that they exist as a result of bad records in the NOAA *ZTD* output. An example of such bad records is shown in Figure 9. As such, it is necessary that an outlier detection and removal mechanism be applied before using the data.
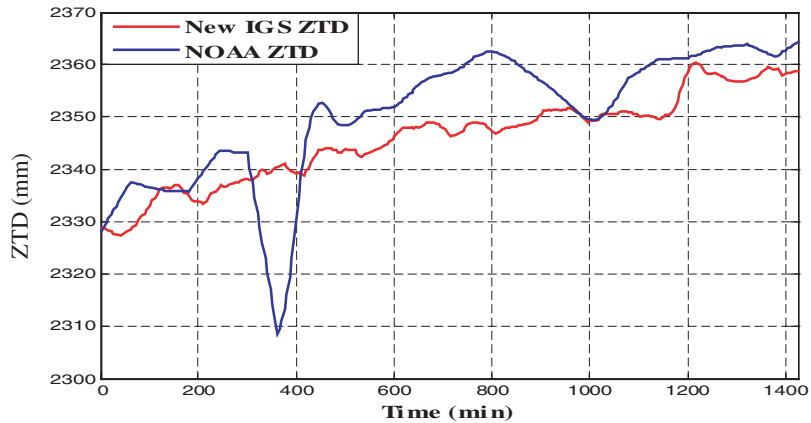
Figure 9. Example of bad records in the NOAA ZTD (day 75 of 2006, USNO station).
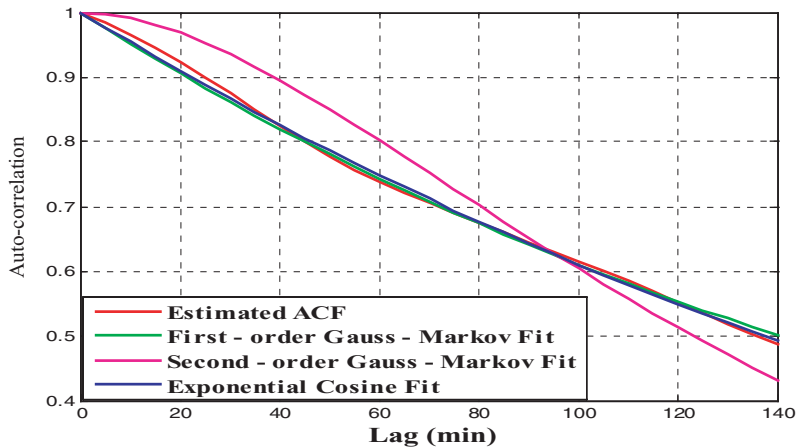

Figure 10. Example of least-squares fit for first-order Gauss-Markov, second-order Gauss-Markov and exponential cosine for NOAA-based residual tropospheric delays (day 215 of 2006, JPLM station).

After data cleaning and linear trend removal, the estimated auto-covariance functions for the NOAA-based daily residual tropospheric delays were obtained using Equation (8). As indicated above, to avoid distortion at large lags, only the first 10% of the estimated ACFs were considered in developing the empirical auto-covariance function. Three functions were tested as potential empirical covariance models, namely the first-order Gauss-Markov, the second-order Gauss-Markov and the exponential cosine models. The coefficients of each of the models were determined through a least-squares fit to the estimated ACFs. Inspecting the RMS of the least-squares fit showed that the exponential cosine function gives the best fit most of the time for all the stations. The results of the first-order Gauss-Markov model were comparable to that of the exponential cosine model. The second-order Gauss-Markov model gave the worst results at all the times. Figure 10 shows an example

Table 1. Mean and standard deviation (STD) of model parameters for NOAA-based residual tropospheric delays.

| Empirical ACF Model	First-Order Gauss-Markov		Second-Order Gauss-Markov		Exponential Cosine 1st parameter		Exponential Cosine 2nd parameter	
Station	Mean	STD	Mean	STD	Mean	STD	Mean	STD
ALGO	0·01036	0·0054	0·02353	0·0095	0·00439	0·0027	0·009328	0·0052
AMC2	0·01510	0·0068	0·03190	0·0117	0·00700	0·0036	0·011600	0·0058
FLIN	0·01110	0·0070	0·02470	0·0121	0·00440	0·0034	0·008300	0·0077
HLFX	0·00780	0·0042	0·01900	0·0074	0·00280	0·0021	0·006700	0·0067
HOLB	0·00840	0·0058	0·01990	0·0101	0·00350	0·0028	0·006000	0·0076
JPLM	0·00897	0·0051	0·02099	0·0091	0·00362	0·0024	0·007373	0·0067
MDO1	0·00924	0·0057	0·02153	0·0101	0·00391	0·0031	0·007428	0·0069
NLIB	0·01170	0·0059	0·02580	0·0102	0·00450	0·0032	0·010000	0·0060
PRDS	0·01584	0·0070	0·03294	0·0119	0·00647	0·0030	0·012625	0·0058
USNO	0·01378	0·0067	0·02943	0·0114	0·00585	0·0030	0·011420	0·0055
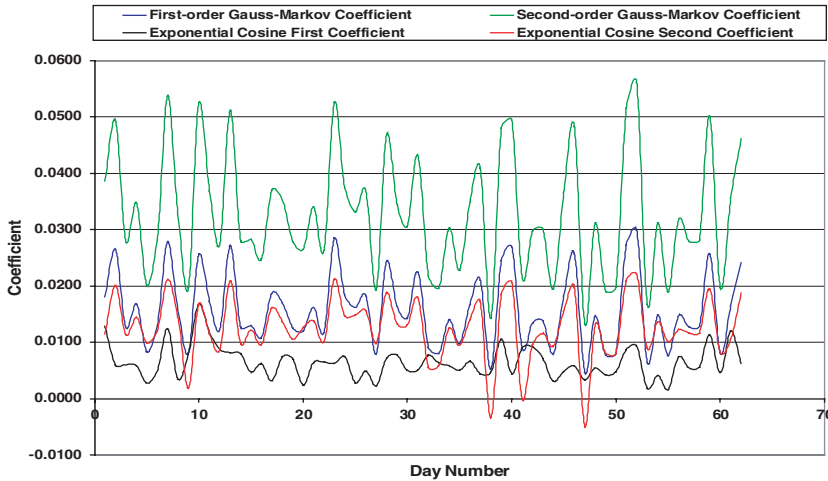
Figure 11. Example of model parameters throughout 2006 for NOAA-based residual tropospheric delays at station PRDS.

of the three models LS fit for day 215 of year 2006 for JPLM station. Figure 11 shows the model parameter variation for each of the three models throughout the year 2006. The first parameter of the exponential cosine model has the least variation, while the parameter of the second-order Markov model has the highest variation. The second parameter of the exponential cosine model and the parameter of the first-order Gauss-Markov model were comparable. Table 1 shows the mean and standard deviation for each of the models' coefficients at all the stations.

An advantage of the first-order Gauss-Markov over other models is its ease of use. Unlike other models, although stochastic modelling with the first-order Gauss-Markov process leads to a fully populated covariance matrix for the observations, its inverse is a block diagonal matrix (El-Rabbany and Kleusberg). This increases

Table 2. Correlation times (minutes) for NOAA-based residual tropospheric delays.

Station	Correlation Time (min)	Station	Correlation Time (min)
AlGO	97	JPLM	111
AMC2	66	MDO1	108
FLIN	119	NLIB	85
HLFX	90	PRDS	63
HOLB	128	USNO	73

the efficiency of the estimation process. As such, without sacrificing the positioning accuracy, the first-order Gauss-Markov model should be adopted to stochastically model the NOAA-based residual tropospheric delay. As there are no significant seasonal variations in the NOAA-based residual tropospheric delay, a single (average) first-order Gauss-Markov model would be adequate for a particular location, regardless of the time of the year. Table 2 shows the average correlation times throughout the year for each station when the first-order Gauss-Markov is used. Examining the correlation times shows that model parameters are affected by the geographical location of the station to a certain degree. In addition, the stations can be clustered into three different groups (or regions), namely eastern, middle and western. The stations in each region have comparable model parameters (correlation times in the case of first-order Gauss-Markov). The average correlation time is in the order of 100 minutes, 75 minutes and 110 minutes for the eastern, middle, and western region, respectively.

5. CONCLUSIONS. Three empirical covariance functions have been tested to stochastically account for the NOAA-based residual tropospheric delays. The exponential cosine model was found to provide the best fit, in the least-squares sense, most of the time for all the stations. The performance of the first-order Gauss-Markov was somewhat comparable. The second-order Gauss-Markov model gave the worst results at all the times. Examining the parameters of the three empirical covariance functions shows that the stochastic characteristics of the NOAA-based residual tropospheric delays are largely season independent. To increase the efficiency of the estimation process without noticeable accuracy degradation, the first-order Gauss-Markov should be used. The NOAA-based residual tropospheric delays can be accounted for stochastically using the first-order Gauss-Markov model with correlation times of 100 minutes, 75 minutes and 110 minutes for the eastern, middle, and western North American regions, respectively.

REFERENCES

Ahn, Y. W., G. Lachapelle, S. Skone, S. Gutman, S. Sahm (2006). Analysis of GPS RTK Performance Using external NOAA Tropospheric Corrections Integrated with a Multiple Reference Station Approach. *GPS Solutions* Vol. 10, No. 3, pp. 171–186.

Byun, S. H., Y. E. Bar-Sever, G. Gendt (2005). The New Tropospheric Product of the International GPS Service. *Proceedings of the ION GNSS 2005*, 14–16 September, Long Beach, CA, pp. 241–249.

El-Rabbany, A. and A. Kleusberg (2004). An Efficient Algorithm for the Inverse of the Fully Populated Covariance Matrix in GPS Relative Positioning. *Journal of Surveying Engineering*, Vol. 130, No. 1, pp. 29–35.

El-Rabbany, A. (1994). *The Effect of Physical Correlations on the Ambiguity Resolution and Accuracy Estimation in GPS differential positioning*. Department of Geodesy and Geomatics Engineering, Technical report No. 170, University of New Brunswick, Canada, 161 pp.

Gao, Y., S. Skone, K. Chen, N. A. Nicholson (2004). Real-Time Sensing Atmospheric Water Vapor Using Precise GPS Orbit and Clock Products. *Proceedings of the ION GNSS 2005*, 21–24 September, Long Beach, CA, pp. 2343–2352.

Gutman, S., T. Fuller-Rowell, D. Robinson (2003). *Using NOAA Atmospheric Models to Improve Ionospheric and Tropospheric Corrections*. U.S. Coast Guard DGPS Symposium, Portsmouth, VA, 19 June.

Héroux, P., J. Kouba, P. Collins, F. Lahaye (2001). GPS Carrier-Phase Point Positiong with Precise Orbit Products. *Proceedings of the International Symposium on Kinematic Systems in Geodesy, Geomatics and Navigation, Banff, Canada*, pp 518–528.

Kouba, J. (2003). *A Guide to Using International GPS Service* (*IGS*). Retrieved January 15, 2007, from http://igscb.jpl.nasa.gov/igscb/resource/pubs/GuidetoUsingIGSProducts.pdf.

Kouba, J., P. Héroux (2001). Precise Point Positioning Using IGS Orbit and Clock Products. *GPS Solutions* vol. 5, no. 2, pp. 12–28.