ZENITH TOTAL DELAY SHORT-TERM STATISTICAL FORECASTS FOR GNSS PRECISE POINT POSITIONING

Karina WILGAN

Wroclaw University of Environmental and Life Sciences, Institute of Geodesy and Geoinformatics, Grunwaldzka 53, 50-357 Wroclaw, Poland

*Corresponding author’s e-mail: karina.wilgan@igig.up.wroc.pl

1. INTRODUCTION

Troposphere delay modeling is one of open problems in real-time positioning, particularly in Precise Point Positioning (PPP) technique. Several troposphere delay empirical models are available as functions of meteorological parameters: temperature, pressure and relative humidity (e.g. Boehm et al., 2007; Leandro et al., 2006). The zenith total delay (ZTD) values are also available as near real-time product of GNSS (Global Navigation Satellite System) processing. To provide fully operational service for real-time PPP (Precise Point Positioning) it is essential to provide real-time ZTD estimates or short-term forecasts from near real-time estimation.

This paper presents statistical approach to predict short-term ZTD from long time series. Several time series models have been used, such as autoregressive model (AR) or autoregressive moving average model (ARMA). Depending on purpose of forecasts, different time series lengths and various prediction horizons have been considered (form 1 to 24 hours). Predictions were included in both global and local model. The global model term means that one statistical model is used for all stations and the local one that each station has its own statistical model. Methods of ZTD prediction have been verified by two independent validators: deterministic Global Pressure and Temperature (GPT2) model and the Numerical Weather Prediction model COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System). The ZTDs were calculated from meteorological parameters and compared with statistical predictions.

This paper presents statistical approach to predict short-term ZTD from long time series. Several time series models have been used, such as autoregressive model (AR) or autoregressive moving average model (ARMA). This approach was used earlier in geosciences studies such as modeling regional rainfall-runoff (Niedzielski, 2007), predicting universal time (UT1–UTC), length of day (LOD) and the axial component of atmospheric angular momentum (Niedzielski and Kosek, 2008) or forecasting sea level anomalies (Niedzielski and Kosek, 2009). Both AR and ARMA models have been proved useful and have been applied in various disciplines such as finances (e.g. Ray, 1993; Porter-Hudak, 1990), transportation (e.g. Lee and Fambro, 1999), medicine (e.g. Ogawa et al., 1993), biotechnology (e.g. Lu et al., 2001) and many others.

Depending on purpose of forecasts, different time series lengths and various prediction horizons have been considered (1-5 hours forecast for positioning, 24 hours forecast for tomography). Predictions are included in both global and local model. The global model term means that one statistical model is used for all stations and the local one that each station has its own statistical model.

This introduction chapter is followed by the data section, which describes the GNSS product used for calculating the statistical forecasts as well as the two external models – Global Pressure and Temperature...
2. DATA

The forecasts were calculated for the ZTD model called IGGHZ-G, calculated on over 100 ASG-EUPOS stations. The forecasts were compared with the model itself as well as with some external models – deterministic global pressure and temperature (GPT2) model and numerical weather prediction model COAMPS (Coupled Ocean/Atmosphere Mesoscale Prediction System).

2.1. IGGHZ-G MODEL

In 2008 Polish Head Office of Geodesy and Cartography established a GBAS called ASG-EUPOS as a part of EUPOS Project (www.eupos.org). At the time of this study, the network consists of 100 Polish and 21 foreign continuously operating GNSS reference stations (Fig. 1). The stations are evenly spread and the mean distance is less than 70 km.

The working group from Institute of Geodesy and Geoinformatics (IGG), Wroclaw University of Environmental and Life Sciences is responsible for modeling and prediction of the state of the troposphere. The ZTD is estimated in half-hour intervals, by processing last 10 hours of GNSS data. A novel processing engine and new Perl scripts were developed for Bernese GPS Software v. 5.0. The ZTD product is called IGGHZ-G (IGG comes from Institute name, H means 1 hour interval, Z is the abbreviation from Zenith and G stands for GNSS) (Bosy et al., 2012). It is run in near-real time regime, 30 minutes after each full hour and use IGS ultra-rapid orbits, clocks and Earth rotation parameters. The GNSS observations are recorded in real-time from RTCM streams, but for some missing stations RINEX files are downloaded as soon as they are available. Processing procedure is adapting the L5/L3 ambiguity resolution strategy (Dach et al., 2007), where wide-lane L5 solution is performed using phase data. Final ZTD estimation is a result from solution, constrained to ITRF2008 coordinates of 23 EPN (EUREF Permanent Network) stations. The a priori troposphere model is Saastamoinen (Saastamoinen, 1972) with dry Niell mapping function (Niell, 1996). Resultant average ZTD’s standard deviation estimated using Bernese GPS Software is ±1.3 mm (min. 0.8 mm, max. 2.0 mm).

2.2. VALIDATION MODELS

The IGGHZ-G ZTD forecasts were validated against ZTD values calculated form meteorological parameters from two external models – GPT2 and COAMPS. The transformation formulas are given in section 3.2.

The empirical GPT (Global Pressure and Temperature) model was created at Vienna University...
of Technology. The model is based on spherical harmonics up to degree and order nine that were adjusted in a least-squares sense to 3 years (September 1999 to August 2002) of 15’ × 15’ global grids of monthly mean profiles for pressure and temperature from the ECMWF (European Centre for Medium-Range Weather Forecasts) reanalysis data (Boehm et al., 2007). The model provides pressure and temperature at any site in the vicinity of the Earth’s surface based only on day of the year, site latitude, longitude and height. Since 2013 the improved model: GPT2 is available (Lagler et al., 2013), which includes also the specific humidity (in this study recalculated into relative humidity). Values of meteorological parameters are given with a 24-hour resolution.

Second validator was the NWP model COAMPS (Hodur, 1997). Mesoscale, non-hydrostatic coupled ocean and atmosphere model COAMPS was built and is being developed by the Naval Research Laboratory in U.S. For this study the model was provided by The Applied Geomatics Centre (CGS) of Military University of Technology in Warsaw (www.cgs.wat.edu.pl). The COAMPS model is based on modules that parameterize important physical phenomena: radiation transport processes, cloud, precipitation, flows in a turbulent boundary layer, moisture flows, vegetation of plants, etc. The description of exact methods of how the COAMPS model is run can be found in Bosy et al. (2010).

The COAMPS outputs are in the form of 3-dimensional matrix, with 30 levels of σ-type vertical coordinate (follows numerical terrain model) and dense horizontal grid (in this study 13x13 km grid was used). Outputs from the COAMPS model that were taken into consideration are: total air pressure, potential temperature calculated into air temperature and water vapor mixing ratio calculated into water vapor partial pressure. The 24-hours predictions with 1-hour resolution are given twice a day; in this study the analyses at 0:00 UTC and 12:00 UTC and the following 11-hours forecasts until the next analyses were used.

3. METHODOLOGY

In this paper 4 models of ZTD were compared: statistical forecasts based on fitting the statistical models into IGGHZ-G data (section 3.1) and ZTD calculated from meteorological parameters from two external models – GPT2 and COAMPS (methods of interpolation and calculation of ZTD are presented in section 3.2).

3.1. STATISTICAL FORECASTS

Two time series statistical models were used: autoregressive model (AR) and autoregressive moving average model (ARMA). To fit a statistical model, the time series must be preprocessed – it cannot exhibit a periodicity or non-stationary. The employed method was to estimate and subtract the deterministic part from the original series (Brockwell and Davis, 1996).

Preprocessed time series \( Y_t \) is considered as an autoregressive AR(r) process, if it is stationary and satisfies for every \( t \) (Brockwell and Davis, 1996):

\[
Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_r Y_{t-r} + Z_t
\]

(1)

where \( (Z_t) \) is a white noise and \( (\alpha_i)_{i=1..r} \) are coefficients, which are estimated using the stepwise least-squares procedure for AR models (Neumaier and Schneider, 2001)). Degree \( r \) is chosen based on the Akaike criterion (AIC), which is \( r \) that minimizes following function (Akaike, 1971):

\[
AIC(r) = \ln(\sigma_r^2) + \frac{2r}{L}
\]

(2)

where \( \sigma_r \) is an estimator of a white noise variance in AR process and \( L \) is the length of time series.

The autoregressive moving average ARMA processes are extensions of AR processes. Time series \( Y_t \) is an ARMA(r,q) process if it is stationary and satisfies for every \( t \) (Brockwell and Davis, 1996):

\[
Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \cdots + \alpha_r Y_{t-r} + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \ldots + \beta_q Z_{t-q}
\]

(3)

where \( (Z_t) \) is a white noise and \( (\alpha_i)_{i=1..r} \) and \( (\beta_j)_{j=1..q} \) are coefficients, which are also estimated using the stepwise least-squares procedure.

3.2. INTERPOLATION METHODS

The COAMPS outputs are provided on a dense grid, but there is a necessity to interpolate meteorological parameters to the locations of GNSS stations. Temperature \( T \) and relative humidity RH were interpolated as a weighted average: (Borkowski et al., 2002):

\[
T = \frac{\sum_{i=1}^{n} T_i \cdot w_i}{\sum_{i=1}^{n} w_i}, \text{ where } w_i = (H - H_i)^{-2}
\]

(4)
K.Wilgan

Fig. 2 Comparison of the original IGGHZ-G time series (blue) with 1-hours statistical predictions of ZTD (ARMA – red, AR – green) and external models (COAMPS – pink, GPT2 – light blue). On the bottom the bias analyses of residuals ZTD\textsubscript{IGGHZ-G-ZTD_{model}}. Data period is 1.12.2013 – 15.03.2013.

\begin{equation}
RH = \frac{\sum_{i=1}^{n} RH_i \cdot w_i}{\sum_{i=1}^{n} w_i}, \text{ where } w_i = \left( (x-x_i)^2 + (y-y_i)^2 + (H-H_i)^2 \right)^{-2}
\end{equation}

(5)

where \( T \), \( RH \), \( H \), \( x \) and \( y \) are temperature, relative humidity, height and horizontal coordinates of \( i \)-th (known) point relatively and \( H \); \( x \); \( y \) are height and horizontal coordinates of interpolated point. For air pressure the method of weighted average was not performing very well, so this parameter was interpolated using modified formula given by Karabatić et al. (2011):

\begin{equation}
p = p_c \left( \frac{T_c - \gamma (H-H_c)}{T_c} \right)^\frac{M}{R}
\end{equation}

(6)

where \( p_c \), \( T_c \) and \( H_c \) are air pressure, temperature and height of the nearest known point, \( \gamma = 0.0065 \text{ [K/m]} \) is a standard temperature lapse rate, \( M = 0.0289644 \text{ [kg/mol]} \) is a molar mass of dry air, \( R = 8.31432 \text{ [N·m/(mol·K)]} \) is an ideal gas constant and \( g \) is a gravitational parameter given by Hitesch (2004):

\begin{equation}
g = 9.8063 \cdot 10^{-3} \cdot \frac{H_c + H}{2} \left( 1 - 0.0026373 \cdot \cos(2\phi) + 5.9 \cdot 10^{-6} \cdot \cos^2(2\phi) \right)
\end{equation}

(7)

where \( \phi \) is the latitude of the interpolated point. All methods of interpolation were tested for the COAMPS model in Wilgan et al. (2015).
4. CASE STUDY

The GNSS ZTD data were collected from 1.12.2012 to 15.03.2013. The results are not representative for all seasons due to data availability. One need to be cautious while extending the study for summer months. The time series of ZTD from IGGHZ-G model were fitted into two statistical models – AR and ARMA. Models were fitted using software Matlab: for degree selection the System Identification Toolbox was used and for model fitting.
Fig. 4 Comparison of the original IGGHZ-G time series (blue) with 24-hours statistical predictions of ZTD (ARMA – red, AR – green) and external models (COAMPS – pink, GPT2 – light blue). On the bottom the statistical analyses of residuals $\text{ZTD}_{\text{IGGHZ-G}} - \text{ZTD}_{\text{model}}$.

Table 1 Average biases and standard deviations of the global model residuals $\text{ZTD}_{\text{IGGHZ-G}} - \text{ZTD}_{\text{model}}$ for all 121 ASG-EUPOS stations.

<table>
<thead>
<tr>
<th>method</th>
<th>absolute bias [mm]</th>
<th>standard deviation [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR(4)</td>
<td>0.60</td>
<td>6.72</td>
</tr>
<tr>
<td>ARMA(8,4)</td>
<td>0.54</td>
<td>6.76</td>
</tr>
<tr>
<td>COAMPS</td>
<td>9.36</td>
<td>7.26</td>
</tr>
</tbody>
</table>
and forecasts functions from package ‘ident’ were utilized. Firstly, all of the time series were preprocessed: a polynomial trend of degree four and periodical trend of degree one were fitted and removed. The residuals of the trend fitting were not found to be white noise, therefore the statistical fitting was considered feasible. The model degree was chosen based on the Akaike criterion, followed by model accuracy validation (in System Identification Toolbox). If there was an insignificant difference in accuracy between low and high model degree, smaller one was chosen, to make model more computational efficient. The predictions were calculated based on the values from the fitted models. The final step was to add the trend again, to make the model comparable to the original IGGHZ-G values. The forecasts were included into both ‘local’ and ‘global’ models – in local model each station was considered separately and has fitted its own model and in the global model every station from the GNSS network has the same statistical model (degree of the model).

4.1. LOCAL MODEL

In local model each one of 121 stations was considered separately and an independent statistical model was fitted. Various prediction horizons have been considered (1-5 hours forecast for positioning, 24 hours forecast for tomography). Figure 2 shows the comparison of the original IGGHZ-G time series for station Katowice (KATO) with 1-hour statistical predictions of ZTD and external models as well as the analyses of the residuals $ZTD_{\text{IGGHZ-G}} - ZTD_{\text{model}}$. The models AR(4) and ARMA(8,4) were found to fit the best to the ZTD data series of station KATO. Models were fitted with one hour resolution (which is the resolution of IGGHZ-G model as well), so the forecast-series is shown as a continuous line. The forecasts start after 120 first hours, which is the first sufficient time-period after which the statistical fitting is consider feasible (Shumway and Stoffer, 2010).

Considering 1-hour forecasts, the statistical models show the best agreement with the IGGHZ-G model. The AR and ARMA models are behaving very similarly, with the residuals of the normal distribution with the mean in 0. The model COAMPS show biases of average 10 mm. The model GPT2 has a 24-hour resolution, therefore it is not following the daily variability of ZTD and in result, the biases of residuals are the highest (distributed from -40 up to 40 mm).

For 1-hour forecasts the statistical models are the best solution. But for PPP purposes longer forecasts are needed. Figure 3 and 4 show 5- and 24-hours predictions (respectively) for station KATO (with the same models degrees). Predictions are shown in 12-hours intervals and only for 10-days for clarity (1.01.2013 – 10.01.2013), but the statistics (biases and standard deviations) were calculated for whole period.

For 5-hours forecasts the residuals $ZTD_{\text{IGGHZ-G}} - ZTD_{\text{model}}$ from statistical models are again evenly distributed with the mean around zero, but with slightly bigger standard deviations (around 5 mm). On the other hand the COAMPS as well as GPT2 model
are behaving in the way as they were with 1-hour forecasts. The 24-hours forecast visibly do not show good agreement (Fig. 4, top), but considering the residuals statistics, the absolute biases are still less than 10 mm and the standard deviations are at the level of 5-10 mm, which is still less than COAMPS/GPT2. Nevertheless, the longer the forecast is, the less accurate. The statistical models cannot predict the longer period (e.g. 24 hours) in the time series variability as they are based on past values of the time series, so if for example the sudden jump occurs, it is impossible that the statistical model detects it. One can conclude that the statistical methods are suitable better for short forecasts (up to 5 hours).

4.2. GLOBAL MODEL

In the global model the same statistical model was fitted for all stations. Models AR(4) and ARMA(8,4) were found to be the most consistent with the global data. Figure 5 shows residuals statistics for all ASG-EUPOS stations. Mean biases and mean standard deviations were calculated for residuals $\text{ZTD}_{i} - \text{ZTD}_{\text{model}}$ from 5-hours statistical forecasts vs. in-situ observations and for COAMPS 5-hours forecasts vs. in-situ observations. Table 1 shows average biases and standard deviations for all 121 stations.

The analyses for local models show that there are only three models worth consideration: both statistical models and NWP model COAMPS. Time resolution of GPT2 model is not good enough to predict daily variability of ZTD so this model was skipped in the global comparisons. The average standard deviations of $\text{ZTD}_{i} - \text{ZTD}_{\text{model}}$ residues for the three remaining models were at similar level – from 6 to 10 mm for most of the country (and average about 7 mm for all models). Although, the biases were definitely smaller for statistical models – with some exceptions on a few sites – there were around 0 (the average absolute biases for all stations are 0.6 mm for AR model and 0.54 mm for ARMA model). For COAMPS model the biases were unevenly distributed across the country – in the center and in the higher mountains at south of Poland the biases were positive and reached even the level of 30 mm; in the rest of Poland there were negative with absolute values 0-10 mm. The average absolute bias value for the whole country is 9.36 mm.

5. CONCLUSIONS

There are several methods to predict future values of zenith total delay. In this paper four of them were presented: empirical GPT2 model, NWP model COAMPS and statistical models: AR, ARMA. The data period was 1.12.2012-15.03.2013, so the comparisons were made only for winter and may not be representable for all seasons. The method with lowest performance was found to be ZTD calculated from meteorological parameters from deterministic GPT2 model (values of residuals biases were spread from -40 mm to 40 mm). The 24-hour time resolution of GPT2 model is not sufficient to predict daily variability of ZTD. The other three methods (statistical models and COAMPS predictions) show similar accuracy, but statistical methods are independent from external data sources such as NWP models. Also, in up to 5-hour forecasts, statistical models show a smaller discrepancy between in-situ observation and a forecast. Accuracy of global model statistical predictions is similar to the accuracy of local models, therefore there is no necessity to perform local predictions. The average absolute biases for 5-hours global forecasts for the entire country are: 0.6 mm for AR model, 0.54 mm for ARMA model and 9.36 mm for COAMPS and average standard deviations are: 6.72 mm for AR model, 6.76 mm for ARMA model and 7.26 mm for COAMPS. The statistical and the COAMPS predictions can be complementary, first one is suitable for shorter (1-5 hours) forecasts and latter for longer (24-hours) forecasts.

ACKNOWLEDGEMENTS

This work has been supported by the Faculty of Environmental Engineering and Geodesy of Wroclaw University of Environmental and Life Sciences (www.aqua.up.wroc.pl), the Ministry of Science and Higher Education: research project N R09 0010 10 (http://www.cgs.wat.edu.pl/ASG_plus) and the Wroclaw Center of Networking and Supercomputing (http://www.wcss.wroc.pl/): computational grant using Matlab Software License No:1019797

REFERENCES


