Methods of detection of changes in the distribution of observed statistics of time derivatives of the total electron content because of cycle slips in navigation satellite’s signals

George Minasyan*, Ivan Nesterov, and Yaroslav Ilyushin

Department of Atmospheric Physics, Moscow State University, Moscow, Russian Federation

Abstract. Based on the analysis of the phase data of the global navigation satellite system, distributions of time derivatives of the L1 phase frequency and the total electronic content are obtained. The change in the distributions of observed statistics of time derivatives of the total electron content was analyzed, because there are cycle slips in signals of navigation satellites. According to the analysis of the statistics of the phase of signals, an assumption about the physical and technical reasons for phase failures was made. The correlation between time derivatives of the phase signals and the total electron content has been obtained, despite the apparent dependence of the latter on the phase of the signal. This ratio showed that neither direct nor inverse dependence of the change in the distribution of time derivatives in both of quantities was found.

1 Introduction

The distance between the GPS satellite and the receiver can be estimated by calculating the pseudorange. To determine the pseudorange or the carrier phase of the radio wave on the GPS satellite – receiver path, the propagation time of the encoded signal should be measured [1]. VHF can produce three pseudo-range code measurements: C/A-code at frequency \(f_1\) (denoted by \(C_1\)), P-code at frequency \(f_2\) (denoted by \(P_1\)) and P-code at frequency \(f_3\) (denoted by \(P_2\)) [2]. The key to the accuracy of the system is the fact that all signal components are precisely controlled by atomic clocks. These highly accurate frequency standards of GPS satellites produce the fundamental frequency of 10.23 MHz. Coherently derived from this frequency are (presently) two signals in the L-band, the \(f_1\) and the \(f_2\) carrier waves generated by multiplying the fundamental frequency by 154 and 120, respectively, yielding \(f_1 = 1572.42\, MHz\) and \(f_2 = 1227.60\, MHz\) [3]. Phase measurements of pseudo-range are calculated with formula: \(D'f = \lambda \Delta \phi + \lambda N\), where \(\lambda = \frac{c}{f}\); \(\Delta \phi\) - measured by the receiver phase difference between the radio wave carrier received from the satellite and the signal of the same frequency generated in the receiver itself. Measured phase pseudo-ranges in radio wavelengths \(\frac{D'f}{\lambda}\) at frequencies \(f_1\) and \(f_2\) are

* Corresponding author: minasyangg.msu@yandex.ru

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).
denoted as $L_1$ and $L_2$, respectively. Full pseudo-ranges along the phase trajectories for the frequencies $f_1$ and $f_2$ can be written as (1):

$$L_1 \lambda_1 = D - \frac{c_4}{f_1} I + \Delta L_0 + \sigma_{\lambda_1}, \quad L_2 \lambda_2 = D - \frac{c_4}{f_2} I + \Delta L_0 + \sigma_{\lambda_2},$$

(1)

$D$ is the true distance between the receiver and the satellite, $\Delta L_0$ is the effective path lengthening due to the influence of the troposphere, clock mismatch, phase ambiguity of measurements, etc., $\sigma_{\lambda_{1,2}}$ - random measurement errors [4]. The value of the total electronic content (TEC) along the beam of the GPS satellite receiver is determined by phase measurements at two frequencies with the formula (2) [5].

$$I = \frac{1}{4 \times 308} \frac{f_1^2 f_2^2}{f_1^2 - f_2^2} \left( (L_1 \lambda_1 - L_2 \lambda_2) + const_{\lambda_{1,2}} + \sigma_L \right),$$

(2)

where $L_1, L_2$ are the pseudo-ranges measured by the phase method in radio wavelengths; $const_{\lambda_{1,2}}$ - ambiguity of phase measurements; $\sigma_L$ - phase measurement error. The generally accepted unit of measurement for TEC is the TECU (Total Electron Content Unit), a unit of measurement of the total electronic content, equal to $10^{16} \left[ \frac{\text{e}}{\text{m}^2} \right]$.

2 Objective

The data from GPS satellites can contain gaps, noise, and various faults that are of a technical nature. In order to use them, pre-processing is necessary. First, because the data is used in TEC calculations. Secondly, the study of cycle slips in signals, their connection with various factors is a separate topic for research considered in this paper.

3 Experimental technique

- Selection of phase data in RINEX files [6].
- Calculation of TEC based on phase data.
- Calculation of the first and the second derivatives for the TEC and for the phase signals $L_1$ and $L_2$.
- Phase data grouping into continuous fragments of phase records for each satellite on signal reception intervals of 15 and 30 seconds.
- Detection of data breaks, as well as the elimination of bad and noisy data sections.
- Classification of gaps that have a physical nature of occurrence and gaps that are not related to them that arise for various technical reasons, for example, malfunctions of equipment.

In the phase data obtained after the pre-processing of the files in the RINEX format, there are various noises, gaps and failures for a number of reasons: technical failure of satellites or stations, external influence on signal generation. We know that the signal is fixed by the station depending on various settings parameters, with different time frequencies: 15, 30, 60 and 120 seconds. Thus, all data was grouped into continuous fragments of phase records for each satellite with a reception frequency of 15 ($\Delta_{15}$) and 30 ($\Delta_{30}$) seconds.

4 RINEX format

The first proposal for the Receiver Independent Exchange. Format RINEX was developed by the Astronomical Institute of the University of Bern for the easy exchange of the Global Positioning System (GPS) data to be collected during the first large European GPS
campaign EUREF 89, which involved more than 60 GPS receivers of 4 different manufacturers [7]. The governing aspect during the development was the following fact: Most geodetic processing software for GPS data use a well-defined set of observables:

- the carrier-phase measurement at one or both carriers (actually being a measurement on the beat frequency between the received carrier of the satellite signal and a receiver generated reference frequency)[7].
- the pseudorange (code) measurement, equivalent to the difference of the time of reception (expressed in the time frame of the receiver) and the time of transmission (expressed in the time frame of the satellite) of a distinct satellite signal.
- the observation time, being the reading of the receiver clock at the instant of validity of the carrier-phase and/or the code measurements [7].

Usually the software assumes that the observation time is valid for both the phase and the code measurements, and for all satellites observed. Consequently all these programs do not need most of the information that is usually stored by the receivers: they need phase, code, and time in the above mentioned definitions, and some station-related information like station name, antenna height, etc. RINEX – hardware and independent format exchange of navigation data. The name of the file is formed as follows – «ssssdddhmm.yy» (Table 1).

### 4 Statistics analysis

Using dual frequency data the first and the second time derivatives were calculated for values \( L_1 \) and the TEC on two time intervals \( A_{15} \) and \( A_{30} \). The statistics of the data from 01.06.2015 (day 151) to 30.06.2015 (day 181) were constructed for all measurements at more than 2000 stations and 32 satellites in one day. For a more detailed study of the behavior of the distribution of the \( L_1 \), and TEC derivatives, histograms were constructed on several analytically significant ranges obtained in the course of calculations.

**Table 1.** RINEX's file name format.

<table>
<thead>
<tr>
<th>symbol</th>
<th>definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>o:</td>
<td>the data file of observations</td>
</tr>
<tr>
<td>yy:</td>
<td>year(two figures)</td>
</tr>
<tr>
<td>mm:</td>
<td>the first minute of observations (00, 15, 30, 45)</td>
</tr>
<tr>
<td>h:</td>
<td>a symbol for ( n^{th} ) hour in day</td>
</tr>
<tr>
<td>ddd:</td>
<td>data file of observations</td>
</tr>
<tr>
<td>ssss:</td>
<td>four-symbolical name of the receiver/antenna</td>
</tr>
</tbody>
</table>

The values of which correspond to the geometry of the satellite's motion relative to the station in the central part are noticeable (Fig.1a and 1b). Since it is known that \( L_1 \lambda_1 \) is the phase path of a radio signal, the satellite speed is a real value, which should vary in a narrow range of values, of the order of \( 4 \times 10^3 \). The central part can be approximated by a parabola. Increasing range construction by several orders, we note the appearance of a multimode structure (Fig.1c and 1d).

Comparing values of derivatives forming this structure with (1) formula, we understand that the derivative with respect to the phase path of the radio signal should not increase or decrease dramatically in a relatively short time \( t = 30 \) sec., since the satellite’s motion relative to the station is uniform, and there can be no abrupt change in the phase path. Consider these values from a physical point of view: suppose that a sharp change caused a rapid increase or decrease in TEC along a satellite-station beam, which should affect the phase derivative without sharp jumps, evenly.
There are histograms of the distribution of the derivative of TEC for 09.06.2015 ($K_p = 0.3$) at various intervals (Fig.2). There are values corresponding to the physical nature of emergence in the central part (hump) (Fig.2a). The hump is approximated by gaussian and power functions with various coefficients. Increasing an interval to one order there are distribution tails which are approximated by power functions (Fig.2b). At increase in an interval of distribution to two orders, we observe the following picture (Fig.2c). Increasing an interval of distribution, promoted emergence of multimode structure of the distribution of tails similar to structure of the distribution for derivative $\frac{dL}{dt}$. Thus, it is possible to make a conclusion that TEC values are formed by various mechanisms of both the physical and technical nature.

Estimation of the TEC value in the ionosphere using the formula (2) explicitly depends on the value $L_1$ and $L_2$. After analyzing constructed histograms of the distribution of the $L_1$ derivative for the considered period of time, the region corresponding to physical values in each of histograms was noticed. Putting the restriction $-3750 < \frac{dL_1}{dt} < 3750$, selecting only the physical part, a new histogram was obtained, similar to the graphs presented above for the derivative of TEC. By constructing this distribution and comparing it with the distribution without taking into account the restriction, we note that this condition did not
significantly affect the distribution of \( \frac{\Delta TEC}{\Delta t} \). Often it happens that for various failures or discontinuities the value \( L_1 \) and \( L_2 \) changes synchronously to a certain constant, and at the same time the TEC derivative does not undergo major change. A similar construction of a TEC with a limit of \(-1 < \frac{\Delta TEC}{\Delta t} < 1\) did not show dependencies for the distribution of \( L_1 \). Significant changes were not noticeable. Therefore, the values of \( \frac{\Delta L_1}{\Delta t} \) and \( \frac{\Delta TEC}{\Delta t} \) are independent of each other.

6 Shares of corrupted data

Failures and ruptures of data can happen not only because of technical malfunction of a system satellite-receiver, but also because of various external factors. To check this assumption, we set restrictions derivant on \( L_2 \) and TEC, which were set above. Quantity of values, which didn't get into the specified interval on each station was calculated. The relationship of faulty values to total is expressed in percentage ratio. The parameter of a share of faulty stations for one day.

For creation of statistics of a share of faulty data on stations parameters of failures (\( P_{\text{fail}} \)) on each station per day during the period from 01.06.2015 (day 151) till 30.06.2015 (day 181) were received. Statistics of failures for time frame \( \Delta 30 \) on day 9.06.2015 (day 159) is given below (Fig.3). On the axis of ordinates the value of parameter of failures as a percentage, and on an abscissa axis the geographic latitude of location of the station is postponed.

7 Conclusion

The conclusions based on this part of the work are as follows:

- The distribution of \( L_1 \) and \( L_2 \) has a complex multimode character.
- The distribution of the TEC in general is smooth, but depending on the histogram interval, the resulting asymptotic distribution with large deviations from the mean is approximated by various power functions.
- The reasons for the appearance of large values in the phase data themselves do not lead to failure of the values of \( \frac{\Delta TEC}{\Delta t} \). Thus, failures are caused by a set of different mechanisms.

![Fig. 3. Statistics of a share of faulty data on stations at 0 < \( P_{\text{fail}} \) < 3% and 3% < \( P_{\text{fail}} \) < 100% at 9.06.2015. - failures of data on a derivative \( \frac{\Delta L_1}{\Delta t} \) and - failures of data on a derivative \( \frac{\Delta TEC}{\Delta t} \).]
The study was partially supported by a grant of the Russian Science Foundation 17-77-20087 “Study of long-period disturbances in the ionosphere, lower atmosphere and ocean by GNSS sounding methods”.

References