Title: Code-Carrier Divergence Monitoring for a Ground Based Augmentation System (GBAS)

Author: Yixie Shao

Advisor: Guillermo González Casado

Department: Applied Mathematics II

Code-Carrier Divergence Monitoring for a Ground Based Augmentation System (GBAS)

Yixie Shao

June 18, 2015
Ground Based Augmentation System (GBAS) is a version of Differential Global Positioning System (DGPS) designed to reliably support airplane’s all phase of approach, landing, departure and surface operations. Local ionospheric gradients are major threats to the GBAS, which prevents the advanced use of GBAS. A new monitor for detecting anomalous ionosphere to provide integrity assurance is implemented in this thesis. The monitor is based on the ”Fast Initial Response” Cumulative Sum (FIR-CUSUM) quality-assurance method, which is sensitive to the small persistent shifts.

The formula of standardized CUSUM in our application is derived and the three parameters of CUSUM, shift in mean $\bar{\mu}_1$, threshold $h$, in-control ”average run length” $ARL$, are introduced. The relationships among these three parameters are studied and the table of these three parameters’ values is calculated in both zero-start and FIR CUSUM based on Markov chain approach. The same threshold $h$ are obtained in both cases for our application. With the derived table, the monitor is implemented by using the data from the nominal days in June, 2013 and January, 2014. The nominal test is carried out for the implementation and the monitor works well as expected.

Potentially, it is useful for the prototype of GBAS implementation by Research group of Astronomy and GEomatics (gAGE) in UPC.
Acknowledgements

First and foremost, I would like to thank my advisor, Professor Guillermo González Casado, who gave me the opportunity to study in this field. During the half year, Professor Guillermo González Casado showed impressive earnestness and patience in his supervision. His valuable suggestions act as a guiding light in the thick mist. Also, I want to give my thanks to all the lecturers and professors of MAMME for their lectures which helps me to build up the skills and knowledge in mathematics. Finally, I want to give my appreciation to my husband, Albert Rosell Anglisano, and my family who support and encourage me always.
# Contents

1 Introduction .................................................. 1
   1.1 Motivation and Main Work ............................ 1
   1.2 Overview of the Ground Based Augmentation System (GBAS) 2
      1.2.1 Global Positioning System (GPS) ............. 2
      1.2.2 Measurement Models ........................... 2
      1.2.3 Error Sources ................................ 4
      1.2.4 Differential Global Positioning System (DGPS) 4
      1.2.5 Ground Based Augmentation System (GBAS) .... 5
   1.3 Ionospheric Delay ...................................... 5
      1.3.1 Previous Work on Ionospheric Delay Gradient .. 5
      1.3.2 Physical Characteristics of Ionosphere ......... 6
      1.3.3 The Influence of Ionosphere on GPS Signal .... 6
      1.3.4 Ionospheric Delay Gradient .................... 7
      1.3.5 Code-carrier Divergence ....................... 7

2 General description of CUSUM .............................. 9
   2.1 Basis for the CUSUM ................................. 9
   2.2 CUSUM Algorithm ...................................... 9
   2.3 Standardized CUSUM .................................. 11
   2.4 Calculating Table of CUSUM Parameters .............. 12
      2.4.1 Transition Matrices ............................ 13
      2.4.2 Calculation of $ARL$ and $h$ .................... 15
      2.4.3 Determination of $\Delta$ ....................... 17
   2.5 $ARL$, $h$ for Zero-Start CUSUM .................... 18

3 Determination of Input Statistic and Parameters for CUSUM
   Algorithm to Monitor Ionospheric Gradients for GBAS 20
   3.1 Smoothed $dz$ As Input .............................. 20
   3.2 True Mean $- \mu_0$ and Standard Deviation $- \sigma$ 22
      3.2.1 In-control Mean $\mu_0$ ......................... 22
      3.2.2 Standard Deviation $\sigma$ .................... 25
   3.3 Out of control Mean $\mu_1$ .......................... 27
   3.4 Threshold $h$ ....................................... 30
CONTENTS

4 Implementation 31
  4.1 Nominal Testing 31
    4.1.1 $h$ and $\sigma$ of Reference Day 31
    4.1.2 Independent Reading as Input 33
    4.1.3 Monitoring Nominal days with FIR-CUSUM 34
    4.1.4 Zero-Start CUSUM 36

5 Conclusions and Future work 38

A Table of CUSUM parameters 43
## List of Figures

1.1 Error sources of GPS measurement .................. 4  
1.2 Example of a GBAS architecture .................... 6  
1.3 Ionospheric delay gradient and GBAS ................. 7  

2.1 Discrete Markov Chain Model for FIR CUSUM ........... 13  
2.2 The curves of in-control ARL and $\bar{\mu}_1$ for given $h$ ........ 16  
2.3 The curve of $\bar{\mu}_1$ and estimated $h$ ............... 17  
2.4 The influence of $\Delta$ on estimated $h$ for given $\bar{\mu}_1 = 0.4$ ...... 18  
2.5 The estimated $h$, Zero-Start CUSUM .................. 19  
2.6 The corresponding in-control ARL for estimated $h$ .......... 19  

3.1 Inflated standard deviation of $dz^s$ and $dz$ ............ 22  
3.2 Normalized out of control mean ($\bar{\mu}_1$) of $dz^s$ and $dz$ .......... 22  
3.3 Histogram of smoothed $dz$ of nominal day-1st, June, 2013 ....... 23  
3.4 $\mu_0$ and $\sigma_{max}$ of smoothed $dz$ of nominal day-1st, June, 2013 .. 23  
3.5 Normalized mean of smoothed $dz$ of nominal day-1st, June, 2013 .... 24  
3.6 Normalized mean of smoothed $dz$ of single satellite ........... 24  
3.7 Standard deviations of three receivers and $\sigma_{max}$ for satellite 22 ... 25  
3.8 Histogram of normalized smoothed $dz$ from Satellite 22 ........... 26  
3.9 PDF of normalized smoothed $dz$ from Satellite 22 ............. 26  
3.10 $\sigma_{inflated}$ of satellite 22 .......................... 27  
3.11 ”true” standard deviation of satellite 22 .................. 27  
3.12 OF1—Equation(3.5); OF2—Equation(3.6) ..................... 29  
3.13 $\bar{\mu}_1$ of satellite 22 ................................. 29  
3.14 Threshold $h$ of satellite 22 ............................. 30  

4.1 $h$ and $\sigma$ of distinct Satellites from day 1st, June, 2013 ....... 31  
4.2 $h$ and $\sigma$ of Satellite 15 from different reference days .......... 32  
4.3 Distinct $h$ and $\bar{\mu}_1$ caused by varying $\dot{I}_{90}$ .................. 33  
4.4 Correlated or Independent $dz$ as input of CUSUM ............... 33  
4.5 130608, Satellite 22, Receiver 2 .......................... 34  
4.6 130608, Satellite 22, Receiver 3 ........................... 35  
4.7 Nominal test results of different satellites ................... 35
LIST OF FIGURES

4.8 Nominal test results of different testing days . . . . . . . . . . 36
4.9 Zero-Start CUSUM, $S(k)$ versus $t$ . . . . . . . . . . . . . . . . . . 36
Chapter 1

Introduction

1.1 Motivation and Main Work

Ground Based Augmentation System (GBAS) is aimed to improve the navigation system’s performance, such as integrity, continuity, accuracy and availability at airport level to support airplane’s all the phase of approach, landing, departure and surface operations. Local ionospheric gradient is an important error source in GBAS. It prevents the advanced use of GBAS. [3] used ”Fast Initial Response” Cumulative Sum (FIR-CUSUM) algorithm to detect small ionospheric gradients and it worked well. However, it is hard to find other papers and publications about application of this methodology in monitoring ionosphere. Thus, this paper focuses on checking the procedure of implementation and testing that proposed in [3] and developing our own monitor based on it. Potentially, it is useful for the prototype of GBAS implementation by Research group of Astronomy and GEomatics (gAGE) in UPC.

The structure of this thesis is the following:

- Chapter 1 is the introduction to GBAS. The measurement models are given here. The error sources of the system are discussed. Especially, the affect of ionosphere in positioning is explained in detail.

- Chapter 2 is the general description of CUSUM algorithm. Both zero-Start and FIR-CUSUM are introduced. The standardized forms of the algorithm – zero-start and ”Fast Initial Response” (FIR) CUSUM – are derived. The relationship among the three parameters of CUSUM – $\hat{\mu}_1$ (equal to the shift in mean for standardized data), $h$ (threshold), in-control $ARL$ (Average Run Length) – is studied and the table of their values is calculated based on Markov chain approach.

- Chapter 3 decides the input (smoothed $dz$) and fixes the parameters (in-control $\mu_0$, out-of-control mean $\mu_1$, standard deviation $\sigma$) of
1.2 Overview of the Ground Based Augmentation System (GBAS)

1.2.1 Global Positioning System (GPS)

The Global Positioning System (GPS) is a satellite-based radio navigation system, which is funded and operated by the United States Department of Defense (DoD). The basic architecture was established in 1973, and the first satellite was launched in 1978. The system was declared operational in 1995. The satellites of GPS lie in medium earth orbit (MEO), which is with altitude from 5,000 to 20,000 km and 2 – 4 orbits per day. Thus, the satellites are typically visible for several hours in each pass over a given location. It provides instantaneous position, velocity and time (PVT) information almost anywhere on Earth at any time and in any weather. The positioning accuracy is 5 m rms in horizontal and 7 m rms in vertical direction [2].

1.2.2 Measurement Models

In GPS system, there are three time scale to deal with: each satellite generates its signals in accordance with a clock on board, i.e. each satellite has its own time $t_s$, and the receiver has its own time $t_u$. Since the satellite and receiver clocks are not synchronized, and each keeps time independently, a common time reference, GPS Time (GPST) $t$, is introduced.

GPS provides two types of measurements: code-phase measurement and carrier-phase measurement. Both of them give the position estimates. Code phase measurement is based on the apparent transit time of the signal from
a satellite to the receiver. While carrier phase measurement is based on an indirect and ambiguous measurement, which does not know the number of cycles of the radio wave transmitted from the GPS satellite. However, once the ambiguity is fixed, the measurement is more precise.

These two types of measurements can be modeled as [2]:

\[
\rho(t) = \rho_{IF}(t) + I(t) + \epsilon_{\rho}(t) \tag{1.1}
\]
\[
\phi(t) = \rho_{IF}(t) - I(t) + \epsilon_{\phi}(t) + \lambda N \tag{1.2}
\]

Where

- \( \rho(t) \) is the code phase measurement, or pseudorange.
- \( \phi(t) \) is the carrier phase measurement.
- \( I(t) \) is ionospheric delay, i.e. delay associated with the transmission of signal through the ionosphere.
- \( \epsilon \) denotes unmodeled effects, measurement errors, such as multi-path error, and receiver noise. The subscripts \( \rho \) and \( \phi \) represent the errors in the code phase and carrier phase measurement, respectively.
- \( N \) integer ambiguity. It cannot be determined in real time. Any break in tracking, no matter how short, could change the integer values. However, if the carrier tracking loop maintains lock, the integers remain constant. It can be solved by using several techniques [4].
- \( \lambda \) is carrier wave length. For GPS \( L_1 \) band, which has turned to be the most important band for navigation purposes with frequency \( f_{L_1} = 1575.42 \ MHz \), \( \lambda_{L_1} = \frac{c}{f_{L_1}} \approx 19 \ cm \). Here, \( c \) is the speed of light, equal to \( 3 \times 10^8 \ m/s \).
- \( \rho_{IF} \) is the ionosphere-free(IF) pseudorange:
  \[
  \rho_{IF}(t) = r(t) + c(\delta t_u - \delta t^s) + T(t)
  \]
- \( r(t) \) is geometric (or, true) range between user and satellite.
- \( \delta t_u \) is the receiver clock bias with respect to GPST \( t \).
- \( \delta t^s \) is the bias in the satellite clock with respect to GPST \( t \).
- \( T(t) \) is tropospheric delay.

Ideally, we would like to measure true range to satellite \( r(t) \). But, in practical, we can only get pseudorange \( \rho(t) \) and \( \phi(t) \), which are biased and noisy measurements of \( r(t) \).
1.2. OVERVIEW OF THE GROUND BASED AUGMENTATION SYSTEM (GBAS)

1.2.3 Error Sources

Code and carrier phase measurements have the same error sources. The errors can be grouped as follows:

- Control segment errors: satellite clock and ephemeris.
- Uncertainties associated with the propagation, including ionospheric delay and tropospheric delay.
- Measurement errors, such as the error due to receiver noise, multi-path etc.

![Error sources of GPS measurement]

Figure 1.1: Error sources of GPS measurement

The errors associated with a satellite clock, ephemeris and the atmospheric propagation are similar for users separated by tens, even hundreds kilometers, and these errors vary "slowly" with time. By taking advantage of this fact, differential GPS (DGPS) is introduced.

1.2.4 Differential Global Positioning System (DGPS)

Differential GPS (DGPS) is an enhancement to GPS that provides improved location accuracy. The basic idea behind DGPS is using some fixed GPS receiver as reference to improve the positioning error of users close to the reference receiver. If the position of the reference receiver is known, the
combined effect of the errors can be estimated for each satellite. Then, the other GPS users in the area can apply these error estimates transmitted by the reference receiver to his own measurement to mitigate the errors and improve the quality of the position estimates. In the implementation of Local-Area differential GPS, the accuracy of positioning can reach meter-level, even sub-meter-level, which depends upon the closeness of the user to a reference station and the delay in the corrections transmitted over a radio link [2].

1.2.5 Ground Based Augmentation System (GBAS)

Ground based augmentation system (GBAS), which is also known as Local Area Augmentation System (LAAS), is a Local-Area differential GPS system. The main goal of GBAS is to provide integrity assurance, additionally, it also increases the accuracy with position errors below 1 m. From a system performance perspective, accuracy is understood as a global system characteristic, whereas integrity is rather intended as real time decision criterion for using or not using the system. The integrity mainly includes following aspects: GBAS transmits augmentation information and associated error bounds to the airplane. Meanwhile, it checks in real time if the error bounds are within protect levels or not. Once the failures are detected, it sends alarms or removes the measurements with failures detected before transmit information.

Figure 1.2 is an example of a GBAS architecture. Augmentation information is generated for each satellite based on observations at ground reference stations at an airport and sent to airplanes for positioning. The service area of GBAS is about 40 km around the airport.

1.3 Ionospheric Delay

1.3.1 Previous Work on Ionospheric Delay Gradient

The ionosphere brings the most worrisome challenge to GBAS system. Many studies about it were carried out. In 1999, [8] obtained statistical bounds for the ionospheric decorrelation effects on a LAAS architecture. The ”moving ionosphere wall” phenomenon was originally discovered with WAAS post-processed and bias-corrected (also knowns as, ”supertruth”) data obtained during ionospheric storms of the recent solar activity maximum [9], [10],[11]. Appropriated threat model of ionosphere based on real data and impact on aircraft positioning accuracy are studied in [7],[9]. Such potentially hazardous ionospheric effects are still not fully investigated or understood yet.
1.3. IONOSPHERIC DELAY

1.3.2 Physical Characteristics of Ionosphere

The ionosphere, a layer of ionized gases, lies approximately 90 km to 1000 km above the Earth. The physical characteristics of ionosphere is highly related with the solar activity and geomagnetic disturbance. The variability of ionosphere is relatively high, i.e. it is diurnal and seasonal, it depends on the phases of the eleven-year solar cycle. Besides, there are also unpredictable short-term effects and localized anomalies (traveling ionospheric disturbances).

The ionosphere is generally well behaved in the temperate zones but can fluctuate near the equator and magnetic poles. The region of highest ionospheric delay is within ±20° of the magnetic equator.

1.3.3 The Influence of Ionosphere on GPS Signal

The speed of propagation of radio signals in the ionosphere depends upon the number of free electrons in the path of a signal. Thus, the ionospheric delay changes according to satellite’s elevation over the horizon. The path length through the ionosphere is shortest in the zenith direction. The zenith delay can vary from several meters to several tens of meters.

In a dispersive medium, two waves of different frequencies travel at different speed. The ionosphere affects GPS signal propagation by delaying code-phase measurements, $\rho$, while advancing carrier-phase measurements, $\phi$. Thus in Equations (1.1) and (1.2), there are opposite signs in front of
1.3. IONOSPHERIC DELAY

the ionospheric delay component, \( I \). Thanks to this character, the delay can be measured from dual-frequency measurements. The GPS navigation data contains an ionospheric model that can remove approximately half of this delay [6].

1.3.4 Ionospheric Delay Gradient

In GBAS system, the augmentation information from the ground reference receivers at airport is used to mitigate errors. However, if there are different error sources between reference and airplane, (as showed in Figure 1.3, the different ionospheric delays in two different paths from a satellite to the reference and to an airplane due to an ionospheric spatial gradient, \( \dot{I} \)), it will greatly affect the performance of augmentation, or even worse, misleads the airplane. In other words, local spatial gradient in ionospheric delay can be an important error source in GBAS.

According to [7], the worst vertical positioning error could be as large as 19 meters due to ionospheric delay gradient. Thus it is important to detect this gradient as quickly as possible.

1.3.5 Code-carrier Divergence

Applying the equations (1.1), (1.2), and replacing \( t \) with \( k \) (here, \( k \) means \( k \)th measurement from a starting epoch \( t_0 \)), then we can get the raw code minus carrier measurement \( z(k) \) as:

\[
z(k) = \rho(k) - \phi(k) = 2I(k) + \epsilon_\rho(k) - \epsilon_\phi(k) - \lambda N
\] (1.3)

Let \( \delta\epsilon(k) \) be the difference of the error term between code-phase and carrier-phase measurement, it can be expressed as \( \delta\epsilon(k) = \epsilon_\rho(k) - \epsilon_\phi(k) \). Using the current measurement \( z(k) \) and the previous one \( z(k-1) \), we get
1.3. IONOSPHERIC DELAY

the following relationship:

\[ dz(k) = \frac{z(k) - z(k-1)}{T} \]

\[ = 2 \frac{I(k) - I(k-1)}{T} + \frac{\delta \epsilon(k) - \delta \epsilon(k-1)}{T} \]  \hspace{1cm} (1.4)

With the estimated ionospheric temporal gradient \( \dot{I}(k) \) as:

\[ \dot{I}(k) = \frac{I(k) - I(k-1)}{T} \]  \hspace{1cm} (1.5)

And \( T \) is the time between two measurements, i.e. the sample interval which is 1 second in our system.

\[ N = S/T \]  \hspace{1cm} (1.6)

- \( S \): the time filter constant, fixed to 100 seconds by GBAS.
- \( N \): the number of samples per second. According to the equation (1.6), it is 100 samples per second in our system.

The divergence rate estimator differentiates the input \( z \) and filters the result using one first order Linear Time-Invariant (LTI) filter to reduce the code noise contributions to the estimation error. The estimator output is \( dv gc \), which is the filtered divergence estimate.

\[ dv gc(k) = \frac{\tau_d - T}{\tau_d} dv gc(k-1) + \frac{T}{\tau_d} dz(k) \]  \hspace{1cm} (1.7)

Here, \( \tau_d \) is the time constant of averaging. In our system, it is set as 100 seconds.
Chapter 2

General description of CUSUM

2.1 Basis for the CUSUM

Cumulative sum (CUSUM) is first proposed in [5]. It can be used to detect small but persistent shifts in the mean or standard deviation of a probability distribution [1]. The method has been widely applied in continuous inspection for quality control. In this thesis, the CUSUM is applied to detect the shift in mean.

To apply CUSUM algorithm, the random process should satisfy three assumptions [1]:

1) The readings are statistically independent and identically distributed.
2) The readings follow a Gaussian distribution.
3) The true mean and standard deviation of this Gaussian distribution are known.

2.2 CUSUM Algorithm

The CUSUM can be divided into one-sided and two-sided CUSUM. In this thesis, we just consider one-sided CUSUM (upward-CUSUM), i.e. the out-of-control mean is bigger than in-control case.

The CUSUM begins at its initial value. In general, the initial value is zero. In this way, the CUSUM is also called zero-start CUSUM. In 1982, Lucas and Crosier suggested a way to improve the performance of the CUSUM for the special situation in which the process mean is already shifted at the time the CUSUM charting begins[1]. It is known as ”Fast Initial Response”
2.2. CUSUM ALGORITHM

(FIR) CUSUM, whose starting value is a positive (nonzero) constant. The "head start" setting as one-half the decision interval is generally accepted as a standard. Thus in this thesis, FIR CUSUM starts at half value of the threshold.

During a run of the CUSUM, if the value of CUSUM is below zero, it is set as zero. And once the value of CUSUM exceeds the threshold, the monitor issues an alarm and needs to be restarted again.

Now, consider a sequence of independent values of a Gaussian random variable, \( x = x(1), \ldots, x(k), \ldots \). The Probability Density Function (PDF) of this random variable is:

\[
P_\mu(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}
\]

(2.1)

Here, \( \mu \) is the mean of the distribution, and \( \sigma \) is the standard deviation.

In the analysis of the CUSUM mean method, a log-likelihood ratio, \( Z(k) \), is first defined [1]:

\[
Z(k) = \ln \frac{P_{\mu_1}(x = x(k))}{P_{\mu_0}(x = x(k))}
\]

(2.2)

Here, \( \mu_0 \) is the mean of in-control case, and \( \mu_1 \) is the mean of out-of-control case, with \( \mu_1 > \mu_0 \).

The key statistical property of this ratio is as follows: a change in the parameter \( \mu \) is reflected as a change in the sign of the mean value of the log-likelihood ratio [12]. In the other words: we can distinguish the real distribution of random variable \( x(k) \) is more like to the in-control or out-of-control case according to the sign of \( Z(k) \). Suppose the random variable \( x(k) \) is still in control. Then it is higher probability for \( x(k) \) to fall into the distribution \( N(\mu_0, \sigma) \). In this case, \( Z(k) \) is negative. In the opposite side, \( Z(k) \) is positive when the real distribution of random variables \( x(k) \) satisfies the out-of-control case.

First, let us consider the FIR CUSUM. The intuitive motivation for the FIR CUSUM is that, if the process is already out of control when the charting begins, then starting the CUSUM part way toward the decision interval will hasten the signal. If, however, the process is not out of control, then it is likely that the CUSUM soon drop back to zero, after which the FIR CUSUM behaves like a conventional zero-start CUSUM. A generic FIR CUSUM algorithm with its recursion is given by:

\[
\begin{align*}
D(1) &= \frac{A}{2}, \\
D(k) &= \max(0, D(k-1) + Z(k)), \quad \forall k > 1
\end{align*}
\]

(2.3)
2.3. STANDARDIZED CUSUM

with the signal of alarm when \( D(k) > A \). Here \( A \) is the threshold corresponding to \( D(k) \).

For FIR CUSUM, the initial value \( D(1) \) is set as one-half of the threshold \( A \). If the reading is in control, the \( Z(k) \) should be negative. Then \( D(k) \) should have a negative shift in general till it reaches zero. If \( (D(k-1) + Z(k)) \) is smaller than zero, \( D(k) \) is set as zero. In opposite, if the reading is out of control, \( Z(k) \) will be positive. Then \( D(k) \) has the positive shift until it reaches the threshold \( A \), and it signals with alarm. Then the CUSUM needs to restart.

Now substituting (2.1) in the log-likelihood ratio, we can get expression of \( Z(k) \) in our case:

\[
Z(k) = \frac{\mu_1 - \mu_0}{\sigma^2} (x(k) - \frac{\mu_1 + \mu_0}{2})
\]  

(2.4)

Since \( \mu_1 - \mu_0 > 0 \), we can rescale (2.3) by dividing \( \frac{\mu_1 - \mu_0}{\sigma^2} \) without changing the property of CUSUM. Then we get the new recursion \( C(k) \):

\[
\begin{align*}
C(1) &= \frac{A \sigma^2}{2(\mu_1 - \mu_0)} \\
C(k) &= \max(0, \frac{D(k) \sigma^2}{\mu_1 - \mu_0}) = \max(0, C(k-1) + x(k) - \frac{\mu_1 + \mu_0}{2}), \ \forall k > 1
\end{align*}
\]

(2.5)

with the signal of alarm when \( C(k) > A \sigma^2 / (\mu_1 - \mu_0) \). For \( C(k) \), the corresponding threshold is \( A \sigma^2 / (\mu_1 - \mu_0) \).

2.3 Standardized CUSUM

Throughout this implementation, we will work with the CUSUM of normalized data, which are obtained by first transforming the process readings by subtracting their in-control mean \( \mu_0 \) and dividing by their standard deviation \( \sigma \).

The normalization can be expressed as:

\[
\begin{align*}
X &\sim \mathcal{N}(\mu_0, \sigma) \Rightarrow \bar{X} \sim \mathcal{N}(0, 1) \\
X &\sim \mathcal{N}(\mu_1, \sigma) \Rightarrow \bar{X} \sim \mathcal{N}(\bar{\mu}_1, 1)
\end{align*}
\]

The general form of normalization is written as:

\[
\begin{align*}
\bar{x} &= \frac{x - \mu_0}{\sigma} \\
\bar{\mu}_1 &= \frac{\mu_1 - \mu_0}{\sigma}
\end{align*}
\]

(2.6)
After the normalization, $C(k)$ is replaced as $S(k)$, and the equations (2.5) are transformed to equation (2.7).

Standardized FIR CUSUM in our case is written as:

$$\begin{align*}
S(1) &= \frac{h_{FIR}}{2} \\
S(k) &= \max(0, S(k-1) + \bar{x}(k) - \bar{\mu}_1)
\end{align*}$$

(2.7)

And it sends an alarm when $S(k) > h_{FIR}$. Here, $h_{FIR}$ is the decision interval for standardized FIR CUSUM.

Since the difference between FIR CUSUM and zero-start CUSUM is only the start point, then the formula for zero-start CUSUM can be simply written as:

$$\begin{align*}
S(1) &= 0 \\
S(k) &= \max(0, S(k-1) + \bar{x}(k) - \bar{\mu}_1)
\end{align*}$$

(2.8)

And it sends an alarm when $S(k) > h_{zero}$. Here, $h_{zero}$ is the threshold for standardized zero-start CUSUM.

From above, we can see $\bar{\mu}_1$ and $h$ are the parameters of CUSUM that we should decide for the implementation. $\bar{\mu}_1$ is equal to the shift in mean for standardized CUSUM and the $h$ is the threshold. When the charts crosses the threshold $h$, this indicates that a shift has occurred and actions will be taken to diagnose the shift. And the CUSUM will then be restarted. The whole sequence going from the starting point to the CUSUM crossing the threshold is called a run. The number of observations from the starting point up to the point at which the threshold is crossed is called the run length. The run length is a random variable, having a mean, a variance and a distribution. The mean of this distribution is called in-control "average run length" (ARL). For a fixed shift in mean $\bar{\mu}_1$, the corresponding $h$ is usually fixed by deciding on the minimum tolerable in-control ARL.

In fact, these three parameters of CUSUM, $\bar{\mu}_1$, $h$ and ARL, are not independent: given $\bar{\mu}_1$ and $h$, we can find the ARL, or, given $\bar{\mu}_1$ and ARL, we can find $h$. And the progress to get the table of these three parameters is introduced in section 2.4.

### 2.4 Calculating Table of CUSUM Parameters

Because of the different start points, zero-start and FIR CUSUM have distinct decision intervals $h$ to maintain the same in-control ARL theoretically. FIR CUSUM has improved initial response. The corresponding price for this better performance is that, in order to maintain the same in-control
2.4. CALCULATING TABLE OF CUSUM PARAMETERS

$ARL$, it is necessary to increase the decision interval $h$ slightly. In this section, we introduce the Markov chain approach to calculate the table of CUSUM parameters. Here, we use the FIR CUSUM as an example. The corresponding $h_{FIR}$ is simply written as $h$ during the calculation.

2.4.1 Transition Matrices

The CUSUM has the Markov chain property: given the first $(k-1)$ values of the CUSUM, the previous values have no effect on the $k$th value of the CUSUM, except the $(k-1)$th one. It can be expressed as:

$$Pr(S(k)|S(1), S(2), ..., S(k-1)) = Pr(S(k)|S(k-1))$$ (2.9)

Here, $Pr$ represents the probability.

We can approximate the continuous state space of the CUSUM with discrete Markov chain model. Figure 2.1 shows the discrete Markov chain model with $h/2$ as initial start. Given an $h$, the range of the CUSUM value from 0 to $h$, which is corresponding to in-control states, is divided into $M$ segments with mesh width $\Delta = \frac{h}{M}$, as showed in Figure 2.1. The resulting $M+1$ states can be denoted as $P_0, P_1, \ldots, P_M$. And the out-of-control state can be labeled as $P_{M+1}$.

![Figure 2.1: Discrete Markov Chain Model for FIR CUSUM](image)
2.4. CALCULATING TABLE OF CUSUM PARAMETERS

The transition probability matrix $p$:

$$p = \begin{pmatrix} p_{00} & \cdots & p_{0(M+1)} \\ \vdots & \ddots & \vdots \\ p_{(M+1)0} & \cdots & p_{(M+1)(M+1)} \end{pmatrix}$$

Its element $p_{i,j}$ means the probability of transition from state $i$ to state $j$ for the value of CUSUM, which can be written as:

$$p_{i,j} = \Pr(S(k) \in P_j \mid S(k-1) \in P_i) \quad (2.10)$$

for $i, j = 0, 1, \ldots, M + 1$.

The progress to derive $p_{i,j}$ can be expressed as follows:

1. Suppose $(k-1)$th value of CUSUM — $S(k-1)$ is in in-control state $i$ ($P_i$, $i = 0, 1, 2, \ldots, M$), i.e. it falls into the range $((i - \frac{1}{2})\Delta, (i + \frac{1}{2})\Delta]$, $i = 0, 1, 2, \ldots, M$. Since $\Delta$ is small, we can use the value of middle point in this interval to represent $(k-1)$th value, i.e. $S(k-1) = i\Delta$.

2. Based on equation (2.7), we consider the $k$th value of CUSUM — $S(k)$.

There are three possibilities:

1) It changes to another in-control state $P_j$, but not the $P_0$ (state 0), i.e.

$$S(k) = [S(k-1) + \bar{x}(k) - \frac{\bar{\mu}_1}{2}] \in ((j - \frac{1}{2})\Delta, (j + \frac{1}{2})\Delta]$$

with $j = 1, 2, \ldots, M$.

2) $S(k)$ is equal to zero.

In this situation, there are two cases lead to this value:

$$[S(k-1) + \bar{x}(k) - \frac{\bar{\mu}_1}{2}] < -\frac{\Delta}{2}$$

Or,

$$[S(k-1) + \bar{x}(k) - \frac{\bar{\mu}_1}{2}] \in ((-\frac{1}{2})\Delta, (\frac{1}{2})\Delta]$$

Then, it leads to:

$$[S(k-1) + \bar{x}(k) - \frac{\bar{\mu}_1}{2}] < \frac{\Delta}{2}$$

3) It falls into out-of-control state $P_{M+1}$, i.e.

$$S(k) = [S(k-1) + \bar{x}(k) - \frac{\bar{\mu}_1}{2}] > h + \frac{\Delta}{2}$$
2.4. Calculating Table of CUSUM Parameters

3. Since we are considering the standardized CUSUM. Then $\bar{x}(k)$ in control satisfies the distribution $\mathcal{N}(0, 1)$. Let $F(\bar{x})$ be the cumulative distribution function (CDF) of standard normal distribution $\mathcal{N}(0, 1)$. It is expressed as:

$$F(\bar{x}) = \frac{1}{2}[1 + erf(\frac{\bar{x}}{\sqrt{2}})]$$  \hspace{1cm} (2.11)

Based on the three possibilities above, we can get the $p_{i,j}$ in each case:

1) $[\bar{x}(k) - \frac{\bar{\mu}_1}{2}] \in ((j - i - \frac{1}{2})\Delta, (j - i + \frac{1}{2})\Delta]$

$$\Rightarrow p_{i,j} = F((j - i + \frac{1}{2})\Delta + \frac{\bar{\mu}_1}{2}) - F((j - i - \frac{1}{2})\Delta + \frac{\bar{\mu}_1}{2})$$

with $i = 0, 1, 2, ..., M$, $j = 0, 1, 2, ..., M$.

2) $[\bar{x}(k) - \frac{\bar{\mu}_1}{2}] < (\frac{1}{2} - i)\Delta$

$$\Rightarrow p_{i,0} = F((\frac{1}{2} - i)\Delta + \frac{\bar{\mu}_1}{2})$$

with $i = 0, 1, 2, ..., M$.

3) $[\bar{x}(k) - \frac{\bar{\mu}_1}{2}] > h + (\frac{1}{2} - i)\Delta$

$$\Rightarrow p_{i,M+2} = 1 - F(h + (\frac{1}{2} - i)\Delta + \frac{\bar{\mu}_1}{2})$$

with $i = 0, 1, 2, ..., M$.

4. Now consider the situation that $(k - 1)$th value of CUSUM is in $P_{M+1}$. It will return to start state with the probability 1. i.e. Once the CUSUM exceeds its decision interval, it will be restarted. It can be expressed as $p_{(M+1)j} = 1$, and $p_{(M+1)j} = 0$, with $j = 0, 1, 2, ..., M, j \neq M$.

2.4.2 Calculation of ARL and $h$

Since the matrix $p$ is derived above, we can find in-control ARL according to the following formula [1]:

$$(I - p_{red})\lambda = 1$$  \hspace{1cm} (2.12)
2.4. CALCULATING TABLE OF CUSUM PARAMETERS

Here, \( p_{\text{red}} \) is the ”reduced” transition matrix. It is \( p \) without the last row and column, i.e. it disregards the transitions to and from state \( M + 1 \). \( I \) is the \((M + 1) \times (M + 1)\) identity matrix. \( 1 \) is a vector of length \( M + 1 \) with all the components equal to 1. \( \lambda \) is the vector of \( ARL \) values for CUSUMs starting in the corresponding States 0,1,\ldots,\( M \). So the middle element in \( \lambda \) is the \( ARL \) value for the CUSUM value starting from \( \frac{h}{2} \).

Figure 2.2 shows the relationship between \( \bar{\mu}_1 \) and the common logarithm of \( ARL \) for different \( h \), with \( \Delta \) is set as 0.125. From the figure we can see: when \( \bar{\mu}_1 \) grows, \( ARL \) also increases. And the relationship between them is exponential. Besides, when \( \bar{\mu}_1 \) is fixed, bigger \( h \) results in greater \( ARL \).

Once in-control \( ARL \) can be derived for fixed \( \bar{\mu}_1 \) and \( h \), we can numerically search for the corresponding threshold \( h \) to fixed \( ARL \) and \( \bar{\mu}_1 \) inversely. First, we need to find a gross interval of \( h \) for a fixed \( \bar{\mu}_1 \). Then, we calculate the corresponding in-control \( ARL \) for each \( h \) in this interval with a step \( \delta h \).

In GBAS, according to Category I, the upper limit of in-control \( ARL \) is \( 10^7 \). Thus the estimated \( h \) is the one whose corresponding in-control \( ARL \) is closest to and below the value of \( 10^7 \). Here, the step \( \delta h \) depends on \( \Delta \). For fixed \( \Delta = 0.125 \), minimal step \( \delta h \) can be set as 0.25, i.e. the precision of \( h \) can reach 0.25. The relationship between \( \delta h \) and \( \Delta \) will be discussed in section 2.4.3.

According to the step above, we can find the table A.1—the set of discrete estimated \( h \) for \( \bar{\mu}_1 \) satisfying \( \bar{\mu}_1 \in [0.2, 1.2] \), \( \delta \bar{\mu}_1 = 0.025 \), and corresponding in-control \( ARL \) which is smaller than but closest to \( 10^7 \).
2.4. CALCULATING TABLE OF CUSUM PARAMETERS

Figure 2.3 is the curve of $\bar{\mu}_1$ and estimated $h$ for FIR CUSUM. It shows that bigger $\bar{\mu}_1$ results in smaller decision interval $h$ for fixed upper limit of in-control ARL.

![Figure 2.3: The curve of $\bar{\mu}_1$ and estimated $h$](image)

2.4.3 Determination of $\Delta$

Throughout the procedure above, we can find the mesh size $\Delta = \frac{h}{M}$ is a parameter that we should decide for the calculation of transition matrix due to the discrete model. As showed in figure 2.1, we use the value of middle point in the interval to represent the value of all the points in this interval. Due to this approximation, the smaller $\Delta$ is, the better accuracy we will get. However, smaller mesh size $\Delta$ means larger number of segments $M$ used to divide the fixed interval $h$. And $M$ decides the computing cost. Since the size of transition matrix $(M + 2) \times (M + 2)$, we can find the relationship between computational cost and $M$ is quadratic. Thus in practical, we cannot choose $\Delta$ infinitely small because of the computing cost.

Since $M$ is the number of segments, it should be an integer. Additionally, we use FIR CUSUM which starts at one-half of the threshold, corresponding to state $\frac{M}{2}$. Thus, $M$ should be an even number, i.e. $M = 2n$, with $n = 1, 2, 3, \ldots$. While from the definition, we have $h = \Delta M$. Then it can be written as:

$$h = \Delta M = 2\Delta n, \quad \forall n = 1, 2, 3, \ldots$$

Based on this relationship, we can find the minimal value of step $\delta h$ is $2\Delta$, i.e. the precision of estimated $h$ is equal to $2\Delta$. For example, for $\Delta = 1$, the precision of $h$ is 2, and for $\Delta = 0.125$, it can reach 0.25.
2.5. ARL, $H$ FOR ZERO-START CUSUM

In this thesis, we set $\Delta = 0.125$, whose corresponding precision of $h$ is $0.25$. From figure 2.3, we can see the $h \in (10, 60]$ in our implementation, compared with the value of $h$, the precision is tolerable. On the other hand, the $M$ is in the range of $(80, 480]$, the computing cost is also acceptable.

Figure 2.4 shows the influence of $\Delta$ on estimated $h$ for given $\bar{\mu}_1$ equal to $0.4$ in our case. The corresponding precisions of $h$ for $\Delta = 0.0625$, $0.125$, $0.25$, $0.5$ are $0.125$, $0.25$, $0.5$, $1$, respectively. This point is fit in the figure. Besides, in this example, $\Delta = 0.0625$ and $\Delta = 0.125$ have the same estimated $h$.

Figure 2.4: The influence of $\Delta$ on estimated $h$ for given $\bar{\mu}_1 = 0.4$

2.5 ARL, $h$ for Zero-Start CUSUM

Follow the steps in the last section, with start value as $0$, we can get the estimated $h_{zero}$ for the same values of $\bar{\mu}_1$, and the corresponding in-control ARL.

Figure 2.5 shows the estimated $h_{zero}$ for $\bar{\mu}_1 \in [0.2, 1.2]$. Compared with figure 2.3, we can see the estimated $h$ are the same in zero-start and FIR CUSUM. Because in our application the upper limit of in-control ARL is very big, which is $10^7$, the influence of the ”head start” is insignificant. Due to the precision of the numerical process, we get the same estimated $h$ in two cases. Based on this result, we choose FIR CUSUM in our implementation.

Figures 2.6 shows the corresponding in-control ARL for estimated $h$ in two cases. Blue ”•” is for zero-start CUSUM, while red ”+” is for FIR-CUSUM. The y-axis shows the logarithm of ARL. From the figure we can
2.5. ARL, $H$ FOR ZERO-START CUSUM

![Figure 2.5: The estimated $h$, Zero-Start CUSUM](image)

see, the zero-start CUSUM has a slightly higher in-control $ARL$ than FIR-CUSUM for fixed $h$. And the difference is not so big to result in different $h$ because of the precision of numerical procedure. All data can be found in table A.1 Appendix A.

![Figure 2.6: The corresponding in-control $ARL$ for estimated $h$](image)
3.1 Smoothed $d z$ As Input

The CUSUM is aimed at detecting mean changes in raw divergence. So the $dz(k)$ is naturally selected as the input. Equation (1.4) is a definition of $dz(k)$. Now we consider a generic form denoted as $dz^s(k)$:

$$dz^s(k) = \frac{z(k) - z(k - k_0)}{k_0 T}$$

$$= \frac{1}{k_0} \left[ \frac{z(k) - z(k - 1)}{T} + \ldots + \frac{z(k - (k_0 - 1)) - z(k - k_0)}{T} \right]$$

$$= \frac{1}{k_0} \sum_{i=0}^{k_0-1} dz(k - i)$$

$$= 2 \dot{I}(k) + \frac{\delta \epsilon(k) - \delta \epsilon(k - k_0)}{T k_0}$$

(3.1)

Here,

$$\dot{I}(k) = \frac{I(k) - I(k - k_0)}{T k_0}$$

(3.2)

Equation (3.1) is a way to smooth the original $dz$, so $dz^s(k)$ can be called smoothed $dz$. 


3.1. SMOOTHED DZ AS INPUT

The quantity of \(\dot{I}(k)\) that we want to monitor is very small, at the level of centimeter per second. Now we should consider the other term, \(\frac{\delta\epsilon(k) - \delta\epsilon(k - k_0)}{Tk_0}\), in equation (3.1).

Since the magnitude of the error term \(\epsilon\) varies "slowly" with time, it means \(\epsilon(k - k_0)\) is similar to \(\epsilon(k - 1)\). Considering the definition of \(\delta\epsilon(k) = \epsilon_\rho(k) - \epsilon_\phi(k)\), we can also get: \(\delta\epsilon(k - k_0)\) is similar to \(\delta\epsilon(k - 1)\). Then term \(\delta\epsilon(k) - \delta\epsilon(k - k_0)\) can be expanded as:

\[
\delta\epsilon(k) - \delta\epsilon(k - k_0) \approx \delta\epsilon(k) - \delta\epsilon(k - 1) = [\epsilon_\rho(k) - \epsilon_\rho(k - 1)] - [\epsilon_\phi(k) - \epsilon_\phi(k - 1)]
\]

(3.3)

As we know, \(\delta\epsilon_\rho(k)\) is at meter level, \(\delta\epsilon_\phi(k)\) is at centimeter level [2]. Then, \(\frac{\delta\epsilon(k) - \delta\epsilon(k - k_0)}{Tk_0}\) is in the order of meter per second. Here, \(T\) is equal to 1 second. Thus, as \(k_0\) increasing, the value of term \(\frac{\Delta\epsilon(k) - \Delta\epsilon(k - k_0)}{Tk_0}\) decreases. \(dz^s(k)\) can have a higher SNR (Signal-to-noise ratio) with a larger \(k_0\). However, a larger \(k_0\) implies a longer monitor lag. Due to this dilemma, we set \(k_0 = 20\), which is equivalent to 20 seconds at the 1 Hz update rate.

In the following, we use the data from satellite 22, day 1st, June, 2013, as an example to show the difference between \(dz^s\) and \(dz\) as input of CUSUM.

Figure 3.1 shows the inflated standard deviation of \(dz^s\) and \(dz\), respectively. The inflated standard deviation \(\sigma_{inflated}\) is derived from the maximum standard deviation \(\sigma_{max}\) of the data received by three reference receivers multiplying the corresponding inflation factors \(f\). In this implementation, Gaussian over-bounding method is used to derive inflation factors for distinct satellites, which will be introduced in section 3.2.2 in detail. From the figure, we can see the \(\sigma_{inflated}\) of \(dz^s\) (with \(k_0 = 20\)) is much smaller than the one of \(dz\) (with \(k_0 = 1\)).

Figure 3.2 shows the \(\bar{\mu}_1\) of \(dz^s\) and \(dz\). We know that \(h\) increases as \(\bar{\mu}_1\) decreasing and the corresponding \(h\) for the \(\bar{\mu}_1 = 0.2\) is 60, according to the table A.1– Table of CUSUM parameters in Appendix A. As showed in figure 3.2b, the \(\bar{\mu}_1\) of the readings–\(dz\) is within 0.04. Then the corresponding decision interval \(h\) is much bigger than 60. The great \(h\) means big error tolerance. It results in the bad performance of the monitor. Thus, we choose the smoothed \(dz\) as the input of our monitor.

In the following thesis, smoothed \(dz\) is written as \(dz\), the superscript \(s\) is omitted.
3.2. TRUE MEAN – $\mu_0$ AND STANDARD DEVIATION – $\sigma$

Since the input of CUSUM is decided, we should get the true mean and standard deviation of the nominal cases as mentioned in Section 2.1 Point 3).

### 3.2.1 In-control Mean $\mu_0$

In this thesis, we get the estimation of the in-control means $\mu_0$ based on the pre-study of the distribution of the smoothed $dz(k)$ in nominal case.

Figure 3.3 is the histogram (using 500 bins) of original data of smoothed $dz$ from the day 1st, June, 2013. From this figure, we can see, almost all the data fall into the interval $[-0.1, 0.1]$. 
3.2. TRUE MEAN – $\mu_0$ AND STANDARD DEVIATION – $\sigma$

Figure 3.3: Histogram of smoothed $dz$ of nominal day-1st, June, 2013

Now we look into the detail of the data by checking the statistic parameters – mean and standard deviation of the same day. Figure 3.4 show the mean and standard deviation as a function of satellite’s elevation.

Figure 3.4a is the means $\mu_0$ of smoothed $dz$ per satellite in every elevation bin (5° per elevation bin). From the figure, we can see the means of all satellites are around zero with random fluctuation in the order of $10^{-3} \text{m/s}$.

Figure 3.4b shows the standard deviations per satellite in every elevation bin. It is the maximum standard deviation among three receivers for each satellite.

Based on the means $\mu_0$, standard deviations $\sigma$ obtained above, plus the inflation factors $f$ obtained by Gaussian over-bounding method, which is
3.2. TRUE MEAN – $\mu_0$ AND STANDARD DEVIATION – $\sigma$

explained in section 3.2.2, we can get normalized in-control means $\bar{\mu}_0$ for all the satellites. It can be written:

$$\bar{\mu}_0 = \frac{\mu_0}{f\sigma_{max}}$$

Figure 3.5 shows the $\bar{\mu}_0$ of all the satellites from day 1st, June, 2013. We can see that most of $\bar{\mu}_0$ are inside the range $[-0.04, 0.04]$. Compared with the range of $\bar{\mu}_1$, $[0.2, 1.2]$, which we will introduce in section 3.3, they are quite small.

Figure 3.5: Normalized mean of smoothed $dz$ of nominal day-1st, June, 2013

Figure 3.6 shows the $\bar{\mu}_0$ of single satellite from the different days. Here, we choose the satellite 15 and 22 as examples. From the figure, we can see the means randomly fluctuate around zero. Thus we can suppose the true in-control mean, $\mu_0$, is zero.

(a) Satellite 15  
(b) Satellite 22

Figure 3.6: Normalized mean of smoothed $dz$ of single satellite
3.2. TRUE MEAN – $\mu_0$ AND STANDARD DEVIATION – $\sigma$

3.2.2 Standard Deviation $\sigma$

The true standard deviation, $\sigma$, of the input $dz(k)$ is also precomputed based on the statistical analysis of nominal data sets. Since the satellites of GPS have different performance, we deal with the data per satellite in this thesis. In this section, we use Satellite 22 as an example.

The process to obtain $\sigma$ per satellite is based on Gaussian over-bounding method. It follows the steps:

1. Choose a nominal day as the reference day. Here, day 1st, June, 2013 is chose. Calculate the standard deviations per elevation bin ($5^\circ$) of three channels corresponding to the satellite 22, separately. A channel means a path from one satellite to a single reference receiver. In our GBAS, there are three reference receivers. Then select the maximum standard deviations per elevation bin $\sigma_{\text{max}}$ among them.

   Figure 3.7 shows the standard deviations of three channels per elevation bin, expressed as "*" in different colors. The zero standard deviation means there is no data or less than 1000 readings in the corresponding bin. And the purple line is the maximum standard deviation $\sigma_{\text{max}}$.

![Figure 3.7: Standard deviations of three receivers and $\sigma_{\text{max}}$ for satellite 22](image)

2. Compute normalized smoothed $dz$ using the $\sigma_{\text{max}}$ and zero mean $\mu_0$.

   The distribution of the normalized data can be gotten from its histogram (using 500 bins).

   Figure 3.8 is the histogram of normalized smoothed $dz$ from single satellite 22 and nominal day 1st, June, 2013. From the plot, we can see that almost all the point are in [-1,1].

25
3.2. TRUE MEAN – $\mu_0$ AND STANDARD DEVIATION – $\sigma$

Figure 3.8: Histogram of normalized smoothed $dz$ from Satellite 22

3. When the number of points in each bin is divided by the total number of points in all bins, the histogram is transformed into the discretized probability density function (PDF), as the blue dotted line in figure 3.9, where the y-axis has a logarithmic scale.
While in the same figure, the red line is the PDF of a zero-mean Gaussian distribution with $\sigma = 1$ and the orange line is the PDF of a zero-mean Gaussian distribution with standard deviation $\sigma = f$ to over-bound the tails of the real data distribution. In other words, at the two tails, the desired Gaussian distribution should have a larger probability than the real data distribution. However, the inflation factor also needs to be minimized during the over-bounding process in order to protect integrity. Here, the inflation factor $f = 1.2$.

Figure 3.9: PDF of normalized smoothed $dz$ from Satellite 22
4. Once we get the inflation factor, inflated standard deviation can be computed according to $\sigma_{\text{inflated}} = f\sigma_{\text{max}}$, as showed in figure 3.10.

![Figure 3.10: $\sigma_{\text{inflated}}$ of satellite 22](image)

5. Finally, by using the linear interpolation between two discrete standard deviations, we can get the "continuous" standard deviations $\sigma_{\text{continuous}}$ as our estimate of true standard deviation in the implementation. In the following, the $\sigma_{\text{continuous}}$ is simply written as $\sigma$. Figure 3.11 shows the true standard deviation $\sigma$ for satellite 22.

![Figure 3.11: "true" standard deviation of satellite 22](image)

### 3.3 Out of control Mean $\mu_1$

Since the CUSUM is used to detect the shift in mean in this implementation, the "target" shift should be determined. The out-of-control mean $\mu_1$ is
3.3. OUT OF CONTROL MEAN $\mu_1$

constructed as:

$$\mu_1(k) = \mu_0(k) + v(k) \quad (3.4)$$

Here, $v(k)$ is the ”target” shift of ionospheric gradients from the nominal case, which is the amount needed to be detected by monitor. It is defined as:

$$v(k) = I_{90} \times OF(k)$$

The value of $I_{90}$ is a fixed value and represents the vertical gradient. The obliquity factor is defined in [2]:

$$OF(k) = (1 - \left(\frac{R_E \cos(\theta(k))}{R_E + h_l}\right)^2)^{-\frac{1}{2}} \quad (3.5)$$

- $\theta(k)$ is the satellite elevation angle.
- $h_l$ is the mean height of the ionospheric shell. It is usually taken in the range of $300 - 400$ km.
- $R_E$ is the average radius of the earth.

The value of $OF$ ranges from one for the zenith direction to about three for elevation angle of $5^\circ$ [2]. The obliquity factor can be simplified as following:

$$OF(k) = 1.0 + 16.0 \times (0.53 - el(k))^3 \quad (3.6)$$

$el(k)$ is the satellite elevation angle at epoch $k$ in units of semi-circles($1$ semi-circle=$180^\circ$).

Figure 3.12 shows the curves of obliquity factor that are obtained from equation (3.5) – the red line, and (3.6) – the blue line. The parameters $h_l$ and $R_E$ in equation (3.5) are set as $350$ m, and $6.371 \times 10^6$ m, respectively. From the figure, we find the difference between the two curves happens mostly when elevation below $10^\circ$, which is the cut-off elevation for GPS. In other words, below this cut-off elevation, the data is discarded from the application since it is not reliable. Two curves tend to be the same as elevation increasing, when the data of GPS is more reliable. Thus, the simplified form of obliquity factor (equation(3.6)) is chose to get the obliquity factor in this thesis.

Since the true in-control mean $\mu_0$ is assumed as zero, then the out-of-control mean $\mu_1$ can be written as:

$$\mu_1(k) = I \times (1.0 + 16.0 \times (0.53 - el(k))^3) \quad (3.7)$$
3.3. OUT OF CONTROL MEAN $\mu_1$

From equation (3.7), we can see that the value of $\dot{I}_{90}$ determines the "target" shift in mean we want to detect. The CUSUM method is designed to detect effectively small ionospheric gradients, which the existing acceleration-ramp-step test fails to promptly detect [3]. Here, we set the "target" $\dot{I}_{90} = 0.0095 \, m/s$.

Applying $\sigma$ derived from reference day for the normalization process – equation (2.6), we can get $\bar{\mu}_1$ for each satellite. Figure 3.13 is the $\bar{\mu}_1$ for satellite 22 based on the data from day 1st, June, 2013.
3.4 Threshold $h$

Once the $\bar{\mu}_1$ is obtained, we can find the corresponding threshold $h$ according to table A.1 – Table of CUSUM parameters in Appendix A, which is obtained in section 2.4. The linear interpolation is taken between two discrete data.

Figure 3.14 shows the threshold $h$ of satellite 22. The threshold $h$ is elevation dependent. As elevation increasing, the $h$ decreases generally.

![Figure 3.14: Threshold $h$ of satellite 22](image)
Chapter 4

Implementation

4.1 Nominal Testing

4.1.1 $h$ and $\sigma$ of Reference Day

The threshold $h$ and standard deviation $\sigma$ per satellite are derived using data from the reference day. Both of these two quantities are elevation-dependent.

Figure 4.1 show the $h$ and $\sigma$ derived from the data of 1st, June, 2013, for Satellite 2 and 15, respectively. Here, $\dot{I}_{90} = 0.0095m/s$. Each satellite has different performance, it results in the distinct $\sigma$ per satellite, as showed in figure 4.1b. The blue line is the $\sigma$ of satellite 2 and the red line is the one of satellite 15. $h$ is derived from the $\bar{\mu}_1$, which is related with $\sigma$ (see Equation (2.6)). Due to this, distinct threshold $h$ can be obtained per satellite. Figure 4.1a shows the $h$ for satellite 2 (blue line) and satellite 15 (red line). Besides, we can find that bigger $\sigma$ results in higher $h$. As the satellite’s elevation increasing, the $\sigma$ decreases generally, and the $h$ also becomes smaller.

![Figure 4.1](image.png)

(a) $h$ of Satellite 2 and 15  
(b) $\sigma$ of Satellite 2 and 15

Figure 4.1: $h$ and $\sigma$ of distinct Satellites from day 1st, June, 2013
4.1. NOMINAL TESTING

Figure 4.2 shows the $h$ and $\sigma$ obtained from the data of satellite 15 from different reference days. Blue line is the day 1st, June, 2013, and the red line is the day 2nd, January, 2014. Figure 4.2b shows that the $\sigma$ of the data set of the same satellite from the different seasons also varies. However, when the elevation is bigger than $20^\circ$, when the data from the satellite is more reliable, the difference is not so significant compared with the difference caused by the distinct satellites. Figure 4.2a shows the resulting $h$ for the two different reference days. Thus, we can choose a nominal day as the reference day to monitor the ionosphere of the other day in real time.

![Image of Figure 4.2](image)

(a) $h$ of different reference days  
(b) $\sigma$ of different reference days

Figure 4.2: $h$ and $\sigma$ of Satellite 15 from different reference days

From Equation(2.6), we can see that another factor of influence in the value of $h$ is $\dot{I}_{90}$, coming from the term $v$. Greater value of $\dot{I}_{90}$ means the bigger shift in mean $\mu$ that we need to monitor. It results in smaller $h$ when in-control $ARL$ is fixed. Figure 4.3 verifies this point.

In the figure 4.3, blue line corresponds to $\dot{I}_{90} = 0.0095$ m/s, while the red line represents $\dot{I}_{90} = 0.0019$ m/s. Figures 4.3b is the $\bar{\mu}_1$ for different $\dot{I}_{90}$. Since the big $\dot{I}_{90}$ is two times of the small one, then the resulting $\bar{\mu}_1$ of big $\dot{I}_{90}$ is also two times of the small one. And figures 4.3a are the $h$ in these two cases. The shape of these two curves are roughly similar, while the magnitude of $h$ for big $\dot{I}_{90}$ is about one-half of the small one.
4.1. NOMINAL TESTING

Since \( h \) and \( \sigma \) per satellite are obtained, they are used to monitor the data of the other day. Now we use it to monitor the data from another nominal day–8th, June, 2013.

As referred in section 2.1, Point 1), the input of CUSUM should be statistically independent. Figure 4.4 shows the difference between correlated and independent data as inputs of CUSUM. Here, \( \dot{I}_{90} \) is set as 0.0095 m/s. The reference day is 1st, June, 2013; while the testing day is 8th, June, 2013.

For smoothed \( dz \), because of the smoothing procedure, the data in \( k_0 \) seconds are correlated with each other. Figure 4.4a is the CUSUM test result of all the data from satellite 22, receiver 1. It shows that even in the nominal situation, there are still many alarms during the day. Because the
4.1. NOMINAL TESTING

inputs are not independent in this case. While, the data per $2k_0$ second, which are independent, are chose as input in figure 4.4b. And the CUSUM works well in this situation.

Based on this, we choose the samples per $2k_0$ second as the input of our monitor.

4.1.3 Monitoring Nominal days with FIR-CUSUM

The threshold is derived from the reference day–1st, June, 2013. $\dot{I}_{90}$ is set as 0.0095 m/s.

Figures 4.5 and 4.6 monitor the data from another nominal day 8th, June, 2013. As expected, all the points are below the corresponding threshold, which is represented by the red dot-line.

Figures 4.5a is the result of channel (2,22). Channel (2,22) means receiver 2 and satellite 22. From this figure, we can see there is a break of data around 11.5 h. After the break, the monitor restarts again. Figure 4.6a is the result of channel (3,22). Even the data are from the same satellite 22, but $S(k)$ is slightly different from figure 4.5a because of the distinct receivers. Figure 4.5b is the corresponding routine of satellite 22 monitored by receiver 2, i.e. the elevation of satellite along with time. While 4.6b is the one monitored by receiver 3. For receiver 2 and 3 monitor the data from the same satellite, they have the same routine. However, the break in figure 4.5b means, the receiver 2 did not work properly around 11.5 h, which produced the break of data flow and the restarting of the CUSUM.

Figure 4.5: 130608, Satellite 22, Receiver 2

(a) $S(k)$ versus $t$  
(b) the curve of satellite elevation
4.1. NOMINAL TESTING

Figure 4.6: 130608, Satellite 22, Receiver 3

Figure 4.7 show the result in monitoring the data from the same day–8th, June, 2013, but different satellites–15 and 2 respectively. It shows, every single satellite is independent with each other, with its own threshold and CUSUM value $S(k)$.

Figure 4.7: Nominal test results of different satellites

Figure 4.8 show the result from the same satellite 15, but different days–9th, June, 2013 and 2nd, January, 2014, respectively. In this case the shapes of threshold are the same, while there is a shift in time while the satellite has different relative positions in different seasons. And the CUSUM value $S(k)$ varies slightly.
4.1. NOMINAL TESTING

4.1.4 Zero-Start CUSUM

The nominal test results of the monitor based on FIR-CUSUM are showed above. In this section, the zero-start CUSUM is applied in order to compare with FIR CUSUM.

Figure 4.9 shows the curve of $S(k)$ versus time based on zero-start CUSUM. The day used for monitoring is 8th, June, 2013. And the reference day is 1st, June, 2013. The data is from channel (1,15). Comparing figure 4.9 with 4.7a, we can see there is the same threshold $h$ of these two monitor based on zero-start and FIR CUSUM, respectively. However, for the nominal test, almost all $S(k)$ are close to zero in using zero-start CUSUM. While, in FIR CUSUM, $S(k)$ starts at $\frac{h}{2}$, generally decreases to zero, then the following part is similar to the case of zero-start CUSUM if there is no break.
4.1. NOMINAL TESTING

in monitoring.

Based on the comparison, we find that FIR CUSUM can detect the ionospheric gradients much faster than zero-start CUSUM, if the ionosphere is anomalous when the monitor starts. Meanwhile, FIR CUSUM has the same performance as zero-start CUSUM in the other part. Thus, we chose FIR CUSUM in our implementation.
Chapter 5

Conclusions and Future work

In this thesis, the statistic properties of code-carrier divergence data from GPS is characterized. The CUSUM algorithm, which is popular in quality control in industry is introduced and the formulas of zero-start and FIR CUSUM are derived. The FIR CUSUM is chose for the implementation of the monitor for detecting in real time the ionospheric gradient, which is small but hazardous error source in GBAS. As a comparison, the zero-start CUSUM is also implemented. Finally the nominal test is carried out with this monitor.

The relationships among the three parameters $\overline{\mu}_1$, $h$, in-control ARL are studied in Chapter 2. For fixed $\overline{\mu}_1$, the larger $h$ leads to bigger in-control ARL. And for the fixed in-control ARL, $h$ decreases as $\overline{\mu}_1$ increasing. Based on discretized Markov Chain model, The values of three parameters are calculated for both FIR-CUSUM(head start is $h$) and zero-start CUSUM, showed in table A.1 in Appendix A. In our case that the upper limit in-control ARL is very large, the estimated $h$ are the same for both CUSUM, while FIR CUSUM has better performance in detecting the shift in mean if the process is already out-of-control when then charting begins. Thus, the FIR CUSUM is chose in our implementation.

Based on the table of CUSUM parameters, the implementation of monitor is carried out in chapter 3 and 4. The true in-control mean is assumed as zero and the true standard deviation is precomputed using the data from single nominal day. The smoothed and independent $dz(k)$ is chose as the input of the monitor. The ”target” $\dot{I}_{90}$ is set as 0.0095m/s for the nominal testing. Nominal tests means monitoring another nominal day. The monitor works as expected.

FIR CUSUM used for monitoring ionosphere is proposed and described in [3]. However, it is hard to find other papers and publication about this
application. Along with little information, in this thesis, we check the procedure and testing that proposed in that doctor thesis [3]. Finally we develop our own monitor. Compared with the one proposed in [3], we have some improvement in following parts:

1. Different formulas for CUSUM algorithm. A different standardized CUSUM formula is derived in this thesis according to [1], which is more straightforward compared with the one proposed in [3].

2. Different ways to determinate true in-control mean. [3] proposed to estimate in-control mean in real time. After studying statistic property of the nominal days’ data and checking the values of in-control mean calculated in the way proposed in [3], we find it is better to assume true in-control mean as zero. It simplifies implementation and it is more reliable.

3. Different formulas for out of control mean. In this thesis, we use a simplified formula for the construction of out of control mean according to [2].

4. In deciding the standard deviation, our calculation is based on the data set per satellite. Because each satellite has its own performance. Meanwhile, we derive distinct inflation factors $f$ for every satellites. However in [3], it is obtained from the data of all channels. As a result, we get the distinct threshold $h$ for single satellite while there is only one threshold for all channels in [3]. Besides, in this thesis, we use smaller elevation bin ($5^\circ$), and very small steps in deriving the table of CUSUM parameters, table A.1, in order to get more data. Then we can use the linear interpolation between two discrete data for obtaining the threshold $h$.

5. Input for the monitor. In this thesis, based on the study of difference in threshold $h$ caused by using $dz$ and smoothed $dz$ as input, we select the smoothed $dz$ ($k_0 = 20$ is the interval for smoothing). And compared with the result of CUSUM for the all smoothed $dz$, which are correlated with each other, we choose the independent one (the data per $2k_0$) as the final input of our monitor during the nominal testing. While, all the data set of smoothed $dz$ are used in [3].

6. In the procedure for deriving transition matrix. [3] proposed that: when the value of CUSUM is smaller than zero, it is set to the starting state for FIR CUSUM (state $P_M$). While, according to [1], the difference between FIR CUSUM and zero-start CUSUM is only at the value of starting point. Thus, when the value of CUSUM is smaller than zero, it should be set to the state $P_0$. 

39
Since the nominal test is carried out in this thesis, one possible future work would be the failure test to check the performance of this monitor. Besides, in this thesis, only two nominal days (one is 1st, June, 2013, the other is 2nd, January, 2014) were chose to show the impact of reference day in determining threshold. In future, more days in distinct seasons and different years can be studied to decide the threshold.
Bibliography


Appendix A

Table of CUSUM parameters

Table A.1 is the table of CUSUM parameters. It contains $\mu_1 \in [0.200, 1.200]$, corresponding estimated $h$ with the precision 0.25, and the real in-control ARL for both FIR and zero-start CUSUM. The limit of in-control ARL is $10^7$. 
<table>
<thead>
<tr>
<th>$\bar{\mu}_1$</th>
<th>$h$</th>
<th>In-control $ARL(x 10^6)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FIR-CUSUM</td>
</tr>
<tr>
<td>0.200</td>
<td>59.75</td>
<td>9.736</td>
</tr>
<tr>
<td>0.225</td>
<td>54.00</td>
<td>9.689</td>
</tr>
<tr>
<td>0.250</td>
<td>49.25</td>
<td>9.520</td>
</tr>
<tr>
<td>0.275</td>
<td>45.50</td>
<td>9.909</td>
</tr>
<tr>
<td>0.300</td>
<td>42.00</td>
<td>9.371</td>
</tr>
<tr>
<td>0.325</td>
<td>39.25</td>
<td>9.625</td>
</tr>
<tr>
<td>0.350</td>
<td>36.75</td>
<td>9.517</td>
</tr>
<tr>
<td>0.375</td>
<td>34.50</td>
<td>9.214</td>
</tr>
<tr>
<td>0.400</td>
<td>32.75</td>
<td>9.822</td>
</tr>
<tr>
<td>0.425</td>
<td>31.00</td>
<td>9.670</td>
</tr>
<tr>
<td>0.450</td>
<td>29.50</td>
<td>9.828</td>
</tr>
<tr>
<td>0.475</td>
<td>28.00</td>
<td>9.325</td>
</tr>
<tr>
<td>0.500</td>
<td>26.75</td>
<td>9.353</td>
</tr>
<tr>
<td>0.525</td>
<td>25.50</td>
<td>8.857</td>
</tr>
<tr>
<td>0.550</td>
<td>24.50</td>
<td>9.081</td>
</tr>
<tr>
<td>0.575</td>
<td>23.50</td>
<td>8.894</td>
</tr>
<tr>
<td>0.600</td>
<td>22.75</td>
<td>9.662</td>
</tr>
<tr>
<td>0.625</td>
<td>21.75</td>
<td>8.679</td>
</tr>
<tr>
<td>0.650</td>
<td>21.00</td>
<td>8.751</td>
</tr>
<tr>
<td>0.675</td>
<td>20.25</td>
<td>8.524</td>
</tr>
<tr>
<td>0.700</td>
<td>19.75</td>
<td>9.552</td>
</tr>
<tr>
<td>0.725</td>
<td>19.00</td>
<td>8.733</td>
</tr>
<tr>
<td>0.750</td>
<td>18.50</td>
<td>9.297</td>
</tr>
<tr>
<td>0.775</td>
<td>18.00</td>
<td>9.674</td>
</tr>
<tr>
<td>0.800</td>
<td>17.50</td>
<td>9.838</td>
</tr>
<tr>
<td>0.825</td>
<td>17.00</td>
<td>9.778</td>
</tr>
<tr>
<td>0.850</td>
<td>16.50</td>
<td>9.495</td>
</tr>
<tr>
<td>0.875</td>
<td>16.00</td>
<td>9.009</td>
</tr>
<tr>
<td>0.900</td>
<td>15.50</td>
<td>8.350</td>
</tr>
<tr>
<td>0.925</td>
<td>15.25</td>
<td>9.524</td>
</tr>
<tr>
<td>0.950</td>
<td>14.75</td>
<td>8.475</td>
</tr>
<tr>
<td>0.975</td>
<td>14.50</td>
<td>9.396</td>
</tr>
<tr>
<td>1.000</td>
<td>14.00</td>
<td>8.024</td>
</tr>
<tr>
<td>1.025</td>
<td>13.75</td>
<td>8.645</td>
</tr>
<tr>
<td>1.050</td>
<td>13.50</td>
<td>9.208</td>
</tr>
<tr>
<td>1.075</td>
<td>13.25</td>
<td>9.697</td>
</tr>
<tr>
<td>1.100</td>
<td>12.75</td>
<td>7.671</td>
</tr>
<tr>
<td>1.125</td>
<td>12.50</td>
<td>7.847</td>
</tr>
<tr>
<td>1.150</td>
<td>12.25</td>
<td>7.934</td>
</tr>
<tr>
<td>1.175</td>
<td>12.00</td>
<td>7.930</td>
</tr>
<tr>
<td>1.200</td>
<td>11.75</td>
<td>7.835</td>
</tr>
</tbody>
</table>

Table A.1: Table of CUSUM Parameters