Geomatics Engineering

Department of Geomatics Engineering

A Superresolution Based Cellular Positioning System Using GPS Time Synchronization

By

Richard Walter Klukas

December, 1997

UNIVERSITY OF CALGARY

Calgary, Alberta, Canada

THE UNIVERSITY OF CALGARY

A Superresolution Based Cellular Positioning System

Using GPS Time Synchronization

by

Richard Walter Klukas

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF GEOMATICS ENGINEERING

CALGARY, ALBERTA

DECEMBER, 1997

ABSTRACT

A system to estimate the horizontal position of cellular telephones operating in the Advanced Mobile Phone Service is developed and its performance evaluated. This system, named Cellocate™, was designed and developed for the application of automatic location of cellular 911 callers. Cellocate™ implements a superresolution based algorithm to estimate, at a number of cell sites, the time of arrival of signals transmitted by a cellular telephone. The time of arrivals are differenced between the cell sites and hyperbolic trilateration is used to estimate the position of the cellular telephone. Time synchronization between the various cell sites is implemented through the use of a GPS receiver, operating in time transfer mode, at each cell site. Least squares is used in the position estimation process to make redundant observations consistent.

The performance of Cellocate™ is evaluated though the use of simulations and field tests. The simulations indicate that for an actual cellular network in Calgary, Alberta, and for typical propagation losses and multipath effects, Cellocate™ is able to estimate the horizontal position of a cellular telephone with an accuracy of 119 m (67%). Operational field tests, in which the Cellocate™ system was installed in four working cell sites, indicate a positioning accuracy between 189 m and 287 m (67%). Positioning accuracy is shown to be a function of received signal SNR, multipath, geometry, and calibration of the system biases. Significant improvement in the precision of the position estimates is achieved by averaging of the results. Given proper system calibration, as well as adequate geometry and signal strength, Cellocate™ will meet the Federal Communications Commission requirement of 125 m (67%) for automatic 911 caller location.

ACKNOWLEDGMENTS

DEDICATION

To my children Belinda, Ethan and Owen.
Three beautiful designs
by the Master Engineer - Almighty God.

TABLE OF CONTENTS

CHAPTER 3

TOA ESTIMATION **23**

CHAPTER 4

POSITION ESTIMATION **41**

LIST OF TABLES

LIST OF FIGURES

xvi

LIST OF SYMBOLS & ABBREVIATIONS

∇_{oi}	marginally detectable error
$\nabla\hat{\delta}_{oi}$	effect of MDE in i^{th} observation on the estimated parameters
α	probability of Type I error
α_l	complex amplitude of the lth sinusoid
a	null space of a matrix
$a(h_m)$	propagation loss correction for vehicular station antenna height
a_k	amplitude of arrival k
A	data matrix or first design matrix
AOA	Angle Of Arrival
AMPS	Advanced Mobile Phone Service
AWGN	Average White Gaussian Noise
β	probability of Type II error
B	second design matrix
B_i	time bias of ith cell site
B_{i1}	difference in time bias between the first and i^{th} cell sites
χ^2	chi squared distribution
c	propagation speed (speed of light)
C	combination of antenna gains and cable losses
$\mathbf{C_{\hat{d}}}$	covariance matrix of parameter corrections
$\mathbf{C}_i$	horizontal coordinates of ith cell site
$\mathbf{C}_i'$	horizontal coordinates of ith cell site for shifted origin
$\mathbf{C_l}$	covariance matrix of the observations
$\mathbf{C_{\hat{l}}}$	covariance matrix of adjusted observations
$\mathbf{C_{\hat{r}}}$	covariance matrix of estimated residuals to observations

$\mathbf{C_x}$	covariance matrix of the parameters
$\mathbf{C_{\hat{x}}}$	covariance matrix of corrected parameters
$C_{xy}(f)$	magnitude-squared coherence
C/A	Coarse Acquisition code
CDMA	Code Division Multiple Access
CPFSK	Continuous Phase Frequency Shift Keying
CRLB	Cramer-Rao Lower Bound
δ_o	non-centrality parameter
$\delta(t)$	Dirac delta function
$\hat{\mathbf{d}}$	estimated corrections to the parameters
ΔT_{ij}	TOA difference between the ith and j^{th} cell sites
$d(\tau)$	Manchester encoded, upsampled and lowpass filtered preamble
dB	decibel
$\mathbf{D}$	signal correlation matrix
$D(z)$	root MUSIC polynomial
DC	Direct Current (used to refer to zero frequency)
DCC	Digital Color Code
DGPS	Differential GPS
DOP	Dilution Of Precision
DRMS	Distance Root Mean Square
DSP	Digital Signal Processor
$\boldsymbol{e}_i$	unmodeled errors
$E[\ ]$	expected value
E911	Enhanced 911
EDOP	East DOP
f_2	phase lag of 2nd path with respect to LOS
$\phi_{xx}(t)$	autocorrelation of signal $x(t)$

$\phi_{yx}(t)$	cross-correlation between signals $y(t)$ and $x(t)$
Φ	correlation matrix
$\Phi_{xx}(f)$	energy density spectrum of signal $x(t)$
$\Phi_{yx}(f)$	energy density spectrum of convolution of signals $y(t)$ and $x(t)$
f	frequency
f_c	carrier frequency
f_d	maximum frequency deviation
f_{data}	data rate
ft	feet
$\mathbf{f(x,l)}$	mathematical model relating parameters to observables
$\Im[\ ]$	Fourier transform
$\Im^{-1}[\ ]$	inverse Fourier transform
F	fade margin
FFT	Fast Fourier Transform
FCC	Federal Communications Commission
FM	Frequency Modulation
g_i	redundancy number of the ith observation
GPS	Global Positioning System
$h(t)$	channel impulse response
h_b	base station antenna height
h_m	vehicular station antenna height
H	Hemitian transpose
$H(f)$	channel frequency response
$H_F(f)$	combined frequency response of filters
H_a	alternative hypothesis
H_o	null hypothesis
HDOP	Horizontal DOP

Hor. Prec.	Horizontal Precision
Hz	Hertz
I	identity matrix
k	vector of Lagrange correlates
kbps	kilobits per second
kHz	kilohertz
km	kilometres
λ	eigenvalue of **R**
l	vector of observables
$\hat{\textbf{l}}$	vector of estimated observables
L	number of sinusoids or propagation loss
LNA	low noise amplifier
LOCA	Location On the Conic Axis
LOS	Line Of Sight
LS	Least Squares
μs	microseconds
m	metres
M	subarray length or FIR filter order
Mbps	megabits per second
Mchip	megachip
MDE	Marginally Detectable Error
MHz	megahertz
MP	Multipath
MUSIC	Multiple Signal Identification and Classification
ν	degrees of freedom
$v(i)$	complex noise value at the ith sampling time
n	number of satellites
$n(i)$	additive white noise at frequency $f(i)$

$n(t)$	additive Gaussian noise
n(0,1)	normal distribution with zero mean and unit standard deviation
ns	nanosecond
N	number of multipath arrivals, data points, frequency domain points, or cell sites; noise floor at 30 kHz
NDOP	North DOP
NE	Northeast
$NFFT$	number of FFT points
NW	Northwest
p_i	geometrical distance between transmit antenna and antenna of site i
$p(\mathbf{x};\theta)$	probability density function of observed sample set $\mathbf{x}$
P	Precision Code
P_t	transmit power
PDF	Probability Density Function
PLL	Phase Lock Loop
PN	Pseudo-Noise
PSAP	Public Safety Answering Point
$\hat{\theta}$	unbiased estimator
θ_k	phase of arrival k
rx	receiver
$\hat{r}_i^*$	i^{th} standardized residual
$\hat{\mathbf{r}}$	estimated residuals to the observations
$\mathbf{r^o}$	vector of ranges from cell sites to approximate coordinates of telephone
R	oversampling factor or propagation distance
$\hat{R}$	range estimator
$\mathbf{R}$	ensemble averaged data correlation matrix

$\hat{\mathbf{R}}$	estimate of data correlation matrix
R_u	third unknown in circular trilateration with bias differences (range from reference site to cellular telephone)
Rx_i	time bias at site i
RECC	Reverse Control Channel
RF	Radio Frequency
RMS	Root Mean Square
RNG	Random Number Generator
σ	standard deviation
σ^2	variance
$\sigma_o{}^2$	a priori reference factor
$\hat{\sigma}_o{}^2$	estimated variance factor
$s[n;\mathbf{q}]$	n^{th} sample of the signal s with parameter θ
s_k	reflection coefficient of the kth reflection
$s(t)$	transmitted signal
$\mathbf{s}$	column of $\mathbf{S}$
$\mathbf{s}(t)$	time scanning vector
$\mathbf{s}(\omega)$	frequency scanning vector
SNR	signal to noise ratio
$S(F)$	Fourier transform of $s(t)$
$\mathbf{S}$	signal matrix
$\hat{S}_{\text{MUSIC}}(z)$	root MUSIC spectrum
SA	Selective Availability
SA_{error}	range error due to SA
SE	Southeast
SVD	Singular Value Decomposition
SW	Southwest

τ_o	propagation time fro m transmitter to receiver
t_k	time delay of arrival k
T	symbol length in seconds
$\mathbf{T}$	vector of TOA measurements
T_B	time bias in the recorded TOT
T_c	chip length
T_c	time of signal transmission
T_i	time of signal arrival at i^{th} cell site
TDMA	Time Division Multiple Access
TDOA	Time Difference Of Arrival
TDOP	Time DOP
TOA	Time Of Arrival
TOA_i	estimated time of arrival at site i
$T\hat{O}A_{MUSIC}$	root MUSIC TOA estimate
TOT	Time Of Transmission
$\mathbf{TOT^o}$	vector in which each entry is the approximate TOT
Txer	Transmitter
UTM	Universal Transverse Mercator
υ	eigenvalue of $\mathbf{SDS}^H$
$\mathbf{v}$	eigenvector of $\mathbf{R}$ or right singular vector of Φ
$\mathbf{V}_N$	matrix consisting of noise eigenvectors of $\mathbf{R}$
$\mathbf{V}_{noise}$	matrix consisting of right singular vectors of Φ
VDOP	Vertical DOP
ω	angular frequency
$w(t)$	white Gaussian noise
w^{RMS}	misclosure RMS value
$\mathbf{w}$	misclosure vector

$\mathbf{w}^0$	misclosure vector evaluated at approximate parameter values
$W(f)$	weighting function
$x(i)$	i^{th} data point
$x(t)$	cellular telephone signal
$\mathbf{x}$	vector of unknown parameters or frequency domain points
$\mathbf{x}^0$	approximate parameter values
$\hat{\mathbf{x}}$	vector of estimated parameters
$X(f)$	frequency domain representation of $x(t)$
(x_c, y_c)	horizontal coordinates of cellular telephone
(x_i, y_i)	horizontal coordinates of ith cell site
$(\hat{x}_i, \hat{y}_i)$	horizontal position estimate of ith grid point
$y(t)$	signal arriving at cell site
$Y(f)$	frequency domain representation of $y(t)$
z	complex variable

CHAPTER 1

INTRODUCTION

1.1 Introduction

The field of cellular telephone positioning has attracted much attention in recent years. Initially, cellular telephone positioning was seen as advantageous for system management purposes. For instance, as system capacity increases in response to consumer demand, cells decrease in geographical size. Knowing the accurate locations of cellular telephones may assist the system in allocating resources effectively.

More recently, however, the focus of cellular telephone positioning has shifted from system needs to human needs. A significant majority of cellular telephone users cite security and safety as the principal reasons for subscribing to cellular telephone services. A recent incident involving a cellular telephone in an emergency situation serves as an excellent example. In February of 1997 a person became stranded along a highway during a winter blizzard (Associated Press, 1997). She used her cellular telephone to call for help but could not provide her location due to white-out conditions. To identify the caller's approximate position, authorities asked her to tell them when she could hear the search plane flying above. From the time of her first call, forty hours elapsed before a ground rescue team reached her. An automatic positioning system would have allowed rescuers to reach her far sooner.

Other security applications also exist. Fraudulent telephone use is of major concern to cellular service providers. Identifying the location of a fraudulent user obviously assists in their apprehension. The same may be said for persons who use cellular telephones for criminal purposes. Other applications include the tracking of stolen vehicles containing cellular telephones as well as important persons carrying cellular telephones.

1.2 Motivation

The most immediate motivation for the development of cellular telephone positioning systems is enhanced 911 (E911) services. In June of 1996, the Federal

Communications Commission (FCC) in the United States created rules that require cellular service providers to provide the position of cellular 911 callers to Public Safety Answering Points (PSAPs). When 911 is called from a land-line telephone, the street address of that telephone is automatically provided to the 911 operator. This service is one of the features of E911. Automatic telephone location allows the 911 operator to dispatch emergency response teams immediately. It is most critical when the caller does not know their location or cannot speak.

Currently, E911 services are not available for wireless subscribers. The FCC would like to change this due to the increasing number of 911 calls made from cellular telephones. The FCC rules state that cellular service providers must provide the horizontal position of a cellular 911 caller with an accuracy of 125 metres in 67% of all cases (FCC, 1996). These rules take effect in October of 2001.

1.3 A Cellular Telephone Positioning System

In response to the FCC mandate, a collaborative effort to design and develop a cellular telephone positioning system was undertaken by Cell-Loc Inc. and the Department of Geomatics Engineering at the University of Calgary. The result of this effort is Cellocate™, a system to estimate the horizontal position of cellular telephones operating under the Advanced Mobile Phone Service (AMPS). At this time, AMPS is the dominant cellular standard in North America and several countries outside North America.

This dissertation is a record of much of the design, development, and analysis of the system to July of 1997. The principal contributions of the author were the incorporation of GPS (Global Positioning System) time synchronization, the positioning algorithms, an analysis of the effect of multipath (MP) on time of arrival (TOA) and position estimation, and testing and analysis of the system. The author also conducted simulations to investigate the performance of the TOA estimation and position estimation processes and determine the system's potential. Cell-Loc Inc. provided the basic TOA estimation

algorithms, the system hardware, and digital signal processing software. Both parties labored through the development stage to produce a working prototype.

The Cellocate™ prototype estimates, at several cell sites, the TOA of a cellular telephone transmission on the reverse control channel (RECC). TOA differences are formed between the cell sites in order to eliminate the time of transmission (TOT). Trilateration is then used to estimate the horizontal position of the cellular telephone given the known positions of the cell sites. GPS is used to synchronize the cell sites in time.

Although most of the concepts on which Cellocate™ is based are not new, such a system has only become realizable in the last year. There are two main reasons for this, the first being TOA estimation accuracy. In the past, TOA estimation by Fourier based correlation has not been of sufficient accuracy to produce a position accuracy of any practical benefit. This has changed with the introduction of superresolution algorithms. As shown by Dumont (1994), one such algorithm, Multiple Signal Identification and Classification (MUSIC), is able to provide TOA estimates of far higher resolution than those produced by Fourier methods. Dumont simulated the use of MUSIC in TOA estimation for the purpose of hyperbolic trilateration. The Cellocate™ system, however, is the first application of MUSIC to AMPS based cellular telephone positioning.

One of the major error sources for TOA estimation of cellular telephone signals is multipath. In the cellular frequency band of 800 - 900 MHz, reflection, refraction and diffraction of the signal, due to man-made and natural objects, is common. As a result, the signal travels from the transmitter to the receiver via many paths. A line of sight (LOS) path may or may not exist. If a LOS path does not exist, the TOA estimate will obviously be later than it should. Since positioning by trilateration assumes LOS propagation, error is introduced. If a LOS signal does exist, it will more than likely be accompanied by later arrivals. If the TOA estimation method is not able to resolve the LOS signal from the others, the TOA estimate will be skewed late, again introducing error. Due to its high resolution, MUSIC is able to provide TOA estimation accuracy, in the presence of multipath, never before realizable.

The second factor which allowed a system such as Cellocate™ to be developed is the implementation of GPS. Since achieving initial operational capability in December of 1993 (Leick, 1995), GPS has been able to provide timing synchronization at the same level of accuracy as atomic clocks. In addition, it has two distinct advantages - GPS receivers are far less expensive than atomic clocks and do not require periodic calibration.

1.4 Research Objectives

The principal objective of this research is to design and develop the Cellocate™ system described above. A second major objective is to investigate and test the various components of Cellocate™ and to analyze the entire integrated system under normal field operation conditions. Since Cellocate™ is designed to address the E911 problem for wireless subscribers, a key objective is to determine whether and under what conditions it is able to meet, in an urban environment, the FCC accuracy requirement of 125 m (67%). Other objectives are to analyze the performance of TOA estimation by MUSIC, investigate the use of GPS for time synchronization, analyze the positioning performance of hyperbolic trilateration, quantify the effect of urban multipath on TOA and position estimation, and to analyze the geometry of the current cellular network in Calgary, Alberta, and determine its effect on positioning accuracy.

These objectives are met through simulations and field tests. Simulations were conducted for TOA estimation in the presence of multipath and position estimation for the Telus Mobility cellular network in Calgary. Field tests were conducted by installing Cellocate™ prototypes in four Telus Mobility cell sites.

1.5 Dissertation Overview

This dissertation consists of nine chapters. Chapter 2 is a discussion of the various approaches to cellular telephone positioning and includes a literature review of past and concurrent research efforts. Chapter 3 deals with TOA estimation theory. The emphasis is on TOA estimation by correlation with Fourier methods and superresolution. Position estimation is discussed in Chapter 4. Least Squares (LS) estimation and various trilateration

models are explained. The chapter ends with Plane Intersection - a closed form position estimation algorithm. Simulation of the multipath corrupted cellular channel is the topic of Chapter 5. The TOA of a signal transmitted through the simulated channel is estimated with MUSIC. Chapter 6 deals with positioning simulations. Chapters 7 and 8 present field test results. A static multipath test with four Cellocate™ prototypes in an open field is discussed in Chapter 7. Results from a test, in which Cellocate™ was installed at four working cell sites, are presented in Chapter 8. In Chapter 9, conclusions and recommendations for further investigations are made.

CHAPTER 2

CELLULAR POSITIONING TECHNOLOGY

2.1 Introduction

      This chapter consists of three main sections. In the first, general methods of cellular positioning are discussed and compared. The second section gives an overview of the work already done in the area of cellular telephone location. Both university and industrial efforts are discussed. The chapter ends with a presentation of the Cellocate™ system, the main object of this dissertation.

2.2 Approaches to Cellular Positioning

      Cellular telephone positioning systems may be divided into two general groups. Those that place the positioning technology in the mobile terminal (that is the cellular telephone) may be classified as mobile terminal centric. Technology within the cellular telephone estimates position and transmits that information to the network for use in applications such as E911. The most obvious example of this type of approach is the inclusion of a GPS receiver in the cellular telephone.

      GPS is a satellite based positioning system designed to provide three dimensional position and velocity information 24 hours a day, in any weather, anywhere in the world (Hofmann-Wellenhof et al., 1993). GPS satellites orbit the earth at approximately 20 000 km and transmit signals which are used by GPS receivers to position themselves. The satellites transmit on two separate frequencies, 1575.42 MHz and 1227.6 MHz. This allows corrections to be made to compensate for signal delay through the ionosphere. On the first frequency (1575.42 MHz) two spread spectrum codes are modulated. The C/A code is a 1023 bit Gold code (Spilker, 1980) with a rate of 1.023 Mbps. The P code, having a period of 38 weeks and a rate of 10.23 Mbps, is also modulated onto the second carrier of 1227.6 MHz. Each satellite has its own unique C/A code and a one week segment of the P code.

The position of the GPS receiver is estimated by spherical trilateration. The TOA of the satellite signal at the receiver is estimated by correlating the spread spectrum codes with replicas in the receiver. The TOT of the signal is part of the information transmitted from the satellite to the receiver. Subtracting this TOT from the estimated TOA and multiplying by the speed of light gives a range from the satellite to the receiver. This range defines a sphere, centered at the satellite, on which the receiver must be located. The position of the satellite is also transmitted to the receiver. If three satellites are used, the three corresponding spheres will intersect at the location of the receiver. In reality, a fourth satellite is required because the offset of the receiver clock, with respect to GPS system time, is also unknown. Selective availability (SA) is an intentional degradation of GPS accuracy usually implemented by dithering the satellite clock correction parameters.

The advantages of a GPS based system are positional accuracy and the availability of the technology. The accuracy of GPS in single point C/A mode is 100 m 2 dRMS (95%) in the horizontal and 156 m (95%) in the vertical component (Lachapelle, 1992). This more than meets the FCC specification. In addition, GPS receivers can be inexpensive and are readily available. However, two serious disadvantages prevent such a system from being a viable option. Currently, there are 50 million cellular telephones in use in North America, none of them containing GPS receivers. To ask consumers to purchase cellular telephones containing GPS receivers, in order to be compatible with a location finding system, is unreasonable. Ideally, any positioning system should be compatible with the telephones currently in use and, as far as possible, transparent to the entire network. For a GPS based system, changes to the signal standards would be required in order to accommodate the position information. Secondly, such a system would suffer the same LOS availability limitations as GPS. Due to a free space path loss in excess of 180 dB, GPS signals require a clear unobstructed path between the receiver and the satellites. GPS is unavailable inside buildings, vehicles, underground parkades, tunnels, etc., areas considered critical from an emergency point of view.

The second group of systems are network centric. In these systems the positioning technology resides in the network rather than the telephone. In this way the telephone need

not be modified and the system will be compatible with telephones currently in use. This is by far the greatest advantage of this approach over the mobile terminal centric approach. Network centric systems rely on radio location techniques. Either or both the angle of arrival (AOA) and TOA of signals transmitted between the cellular telephone and network cell sites are estimated. Position may then be estimated by triangulation in the case of AOA or trilateration in the case of TOA. Triangulation refers to the intersection of bearings from at least two stations with known location. An example of triangulation is given in Figure 2.1.
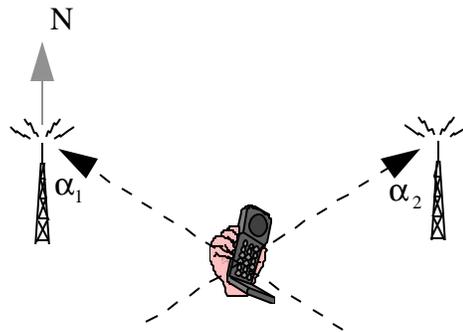
Figure 2.1   Location by Triangulation

In Figure 2.1, α_1 and α_2 refer to AOA estimates of a signal transmitted by the cellular telephone and received at cell sites 1 and 2 respectively. The AOA estimates are obviously with respect to some reference direction such as North. The position estimate is then the intersection of the bearings from the two cell sites.

Trilateration involves the intersection of position lines formed by ranges or range differences. Consider the scenario of Figure 2.2. Pictured are three cell sites receiving signal transmissions from the cellular telephone. At each cell site the TOA of the signal is estimated. If the TOT of the signal is known, the range from cell site to cellular telephone may be determined by multiplying the propagation time by the propagation velocity. According to that cell site then, the cellular telephone lies somewhere on a circle (for the

horizontal case) with a radius equal to the estimated range and centered at the cell site. Unless the cellular telephone is exactly midway between two cell sites, the range circles of two cell sites will intersect at two points. Depending on the size of the range circles, the application at hand, and any a priori positional information, it may or may not be obvious which of the two intersections corresponds to the position estimate. For the cellular application, the distances involved require at least a third range, as in Figure 2.2, to identify the position sought. Position estimation by the intersection of circles is herein referred to as circular trilateration.
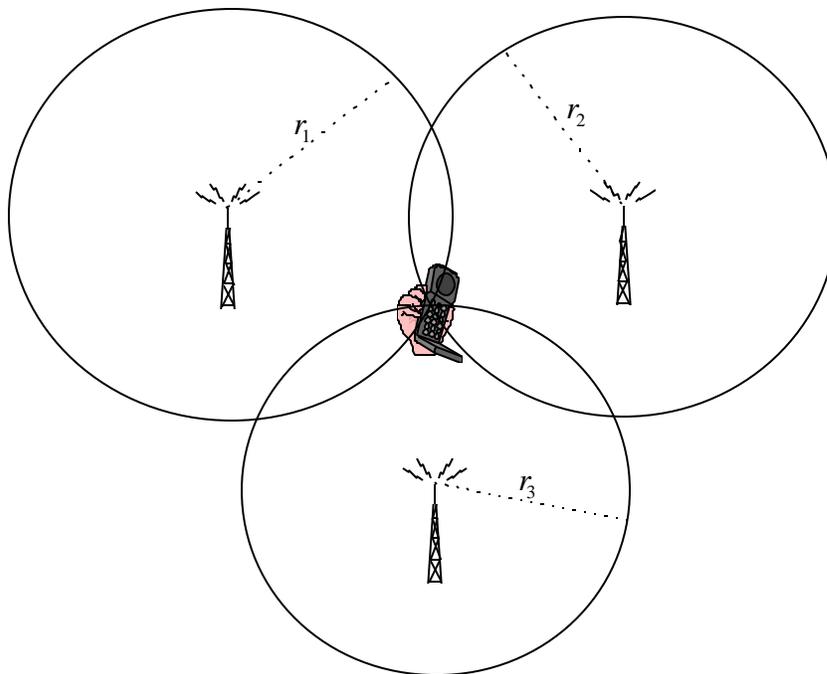
Figure 2.2   Circular Trilateration

In a network centric cellular positioning system, TOT information from the cellular telephone is unavailable. In such a case, the unknown TOT may be eliminated by forming TOA differences between cell sites. Lines of position are then defined as those for which the TDOA (time difference of arrival) between two cell sites is a constant. For the two

dimensional case, the conic section which is defined by this very constraint is the hyperbola (Swokowski, 1983). The relevant half of the hyperbola is known from the arrival order of the two cell sites involved. Two lines of position requires two hyperbolas and therefore three cell sites. The intersection of the two hyperbolas then estimates the position as illustrated in Figure 2.3. The use of TDOAs in this way is termed hyperbolic trilateration.

$$TOA_1 - TOA_3 \qquad\qquad TOA_2 - TOA_3$$
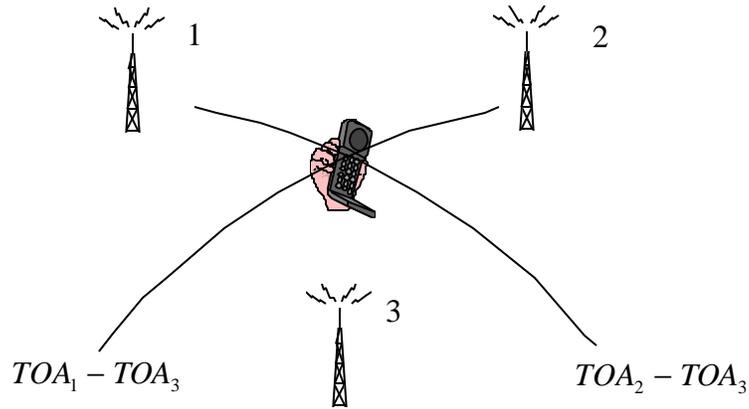
Figure 2.3   Hyperbolic Trilateration

In the past, AOA and TOA estimation accuracy has been the greatest disadvantage of these systems. The accuracy of the derived position is, of course, a function of the accuracy of the AOA or TOA estimates. Accurate estimation of AOA usually requires the use of expensive steered beam antenna arrays and is greatly affected by multipath. AOA estimation using superresolution and virtual antenna arrays has been investigated for the AMPS standard (Klukas, 1993). For a LOS signal in the presence of multipath, an AOA accuracy of approximately $5°$ was possible. This translates into a positional error of approximately 100 m or less for cell sizes on the order of 1 km.

TOA estimation may be accomplished by correlating a received signal with a replica of that signal. The accuracy of TOA estimation by correlation is a function of the signal bandwidth. Spread spectrum signals are, therefore, ideal for this in that they generally

have a relatively large bandwidth. GPS is an example of this type of radio location technology. The C/A code has a clock or chip rate of 1.023 Mbps (Spilker, 1980). The individual elements of the code are referred to as chips. The chip length, being the speed of light divided by the chip rate of 1.023 Mbps, is approximately 293 m. For correlation by delay lock loop, errors due to multipath are limited by the chip length. Multipath reflections arriving later than one and a half chips after the first arrival introduce no error (Van Nee, 1991). Multipath errors are proportional to the early-late spacing of the delay lock loop and reach a maximum of half the early-late spacing (Van Nee, 1992). Early-late spacings of 0.1 chips are now common in GPS receivers (Van Dierendonck et al., 1992).

Correlation of spread spectrum signals is often achieved with traditional Fourier based techniques. In that case, resolution is generally limited by the Rayleigh resolution criterion (Haykin, 1991). This states that two signals are just resolved when the peak of one sits at the first minimum of the other. Since the autocorrelation function of pseudo-random spread spectrum codes is a triangle with a base of twice the chip length, Fourier based correlation techniques are only able to resolve multipath arrivals with delays greater than the chip interval. This corresponds to the limits of the delay lock loop discussed above.

In any case, the smaller the chip interval, the better the multipath rejection. In the Cellocate™ system presented herein, the signal which is used for correlation is the cellular telephone registration preamble transmitted on the reverse control channel. The data rate of this signal is only 10 kbps giving a chip length of 30 km. The resolution of typical Fourier based correlation techniques will obviously not be adequate. Instead, superresolution algorithms may be employed in the correlation process. They are so named because of their ability to exceed the Rayleigh resolution bound. Dumont (1994) has investigated the use of superresolution in spread spectrum ranging systems. Dumont uses root MUSIC in the TOA estimation process and demonstrates the improvement over traditional Fourier correlation. She also demonstrates MUSIC's ability to resolve multipath arrivals within one chip interval. For a 10 Mbps pseudo-random signal, a signal to noise ratio (SNR) of 10 dB, and a two ray channel with ray separation of $1/8$ of a chip, MUSIC was able to resolve both arrivals in 75% of all cases. This same method is used in the Cellocate™ system to estimate

TOA of the 10 kbps registration message preamble. It will be shown with field data that MUSIC in the correlation process outperforms Fourier based correlation in terms of position accuracy.

TOA estimation accuracy in the past would have been inadequate for meeting the FCC specification. However, superresolution has taken TOA estimation to a new accuracy level. This use of superresolution in TOA estimation makes a network centric solution based on trilateration a feasible, low-cost and attractive solution to the cellular telephone positioning problem. TOA estimation by MUSIC will be further discussed in Chapter 3.

2.3 Previous Work

The positioning of cellular telephones has been of interest for some time. It may be considered a special case of the broader field of automatic vehicle location. One means of automatic vehicle location is position estimation of a transmitter located in the vehicle. This is very similar to the application of cellular telephone positioning in that the transmitting device is now the cellular telephone.

Much work was done in the area of vehicle location in the early 1970s. Turin et al. (1972a) performed experiments in the San Francisco area to generate a statistical model of urban multipath propagation. From a fixed site, 100 ns pulses were simultaneously transmitted at frequencies of 488, 1280, and 2920 MHz and received at a mobile van. The data was analyzed and from it statistical models derived. The models were then used in the simulation of an urban vehicle monitoring system and the results presented in a companion paper (Turin et al., 1972b). Two methods of ranging were simulated. In the first, TOA was estimated by measuring the phase of a narrowband (25 kHz) phase modulated signal. In the second, a wideband (10 Mbps) pulsed waveform was employed. TOA was estimated by observing when the envelope of the received signal exceeded a threshold. Simulations of phase-ranging for a dense urban environment resulted in a mean range error of 440 m and a standard deviation of 627 m. For pulse-ranging the corresponding values were 95 m and 65 m.

In either case, the position of a simulated transmitter was estimated by trilateration using TOA estimates from a number of fixed sites. LS was used to fit a position to the TOA observations. Cities with concentric circular propagation environments were simulated with 4, 6, or 8 receivers placed symmetrically along a circle of radius 3048 m. Simulated transmitter points were placed along circles of radii 762 m, 2286 m, and 4572 m. The resulting radial location error distribution for phase ranging and 8 receivers had a mean of 141 m and a standard deviation of 110 m. For a pulse-ranging system the mean error was 43 m with standard deviation 31 m.

Warren et al. (1972) built and tested a vehicle location system based on TOA measurements and hyperbolic trilateration. The vehicle to be positioned emitted a signal consisting of a 3150 Hz tone FM modulated onto a 154 MHz carrier. The phase of this signal, received at each of the monitoring stations, was correlated with synchronization pulses transmitted from a central site. This formed the basis of the TOA measurements. The system was tested with four monitoring stations and the horizontal location accuracy was found to be 549 m (67%).

In recent years the area of vehicle location, and more specifically cellular telephone location, has again attracted a large amount of interest. Spread spectrum based systems, in particular, have been heavily investigated. In the late 1980s, an automatic vehicle location system was developed and put into use in Australia (Hurst, 1989). This system, named QUIKTRAK, measures TOA of a 1 Mchip/sec spread spectrum signal and estimates position by hyperbolic trilateration. The application is fleet management of commercial vehicles. The system is claimed to have an accuracy of approximately 30 m in a coverage area of some 2000 square kilometres. Goud (1991) also analyzed the use of spread spectrum trilateration and reports a range error mean of 18 m and standard deviation of 40 m for a 10 MHz bandwidth. His positioning simulations suggest a mean absolute location error of 26 m and a 90th percentile of 51 m for 8 sensors located on a circle of radius 3 km. Dumont (1994) also simulated positioning accuracy using root MUSIC to estimate TOA of a 10 Mchip/sec spread spectrum signal. With four base stations optimally spaced on a sphere of

radius 200 m, a mobile anywhere within the sphere could be located with a mean error of less than 10 m.

The way in which TOA estimates are processed in order to estimate position has also been investigated. Morley (1995) compares LS with maximum likelihood estimation based on statistical distributions for multipath and iterative hill-climbing algorithms. His simulations show that under certain conditions, maximum likelihood has a 90th percentile position error 1.4 to 3.6 times less than that of LS. Schmidt (1996) and others (Smith et al., 1987) have developed non-iterative LS based algorithms to estimate position from range or TOA differences.

Although much of the attention has focused on TOA technologies, AOA has not been abandoned. Sakagami et al. (1992) investigated the use of multibeam antennas for vehicle positioning in multipath environments. Field tests performed in Tokyo yielded RMS position errors of 200 m or more when measuring AOA at three sites. Sakagami et al. report errors in excess of 1000 m due to strong multipath from distant reflectors.

A number of companies are developing cellular telephone positioning systems. Lockheed Sanders, Inc., offers a two stage positioning system (Sanders, 1996). The first stage, called Smart Look™, provides a wide area (or approximate) location of an AMPS cellular caller. Smart Look™ is installed in the existing cellular infrastructure and requires no modification of the cellular telephone. From dense urban environment field trials, Sanders claims a Smart Look™ location accuracy of just over 3000 feet or 914 m. A handheld device called Micro Look™ may then be used to lead the user directly to the cellular telephone by providing a true line of bearing from signal strength measurements (Microwaves & RF, 1995). Sanders claims a position accuracy better than 125 m for the two stage system. Both Smart Look™ and Micro Look™ use a proprietary technology called "compressive TDOA". Sanders plans to make this system commercially available when market conditions improve.

The Associated Group, Inc. (1997) has developed a product called TruePosition™ which also uses TDOA technology. TruePosition™ is very similar to Cellocate™. Both systems estimate TDOA of transmissions on the reverse control channel, use GPS to provide

timing information, and require no modification of the cellular telephone. In the TruePosition™ system, TDOAs are obtained by cross-correlating between signals received at different cell sites (Stilp et al., 1994). The cross-correlation process is a crucial difference between the two systems. TruePosition™ uses shift registers to perform a sliding correlation whereas Cellocate™ employs MUSIC. The location algorithm in TruePosition™ is a two step process that first finds the minimum LS difference between the measured TDOAs and theoretical values for a grid of hypothetical locations. The best theoretical location is then used as the starting point for a second search with finer resolution. The positioning algorithm in Cellocate™ follows the classical implementation of LS for trilateration as outlined in Chapter 4. In TruePosition™, a quality factor is associated with each TDOA to account for multipath effects. This quality factor is based on the width of the correlation peak which represents the degree of multipath corruption in the TDOA measurement. TruePosition™ is reported to provide a location accuracy of 400 feet (122 m) or better. With a wideband digital system, Associated expects to achieve an RMS error of 114 m.

A TDOA based system has also been developed by Engineering Research Associates. This system also makes use of signals transmitted on the reverse control channel and uses GPS to provide time synchronization (Kennedy et al., 1994). Its uniqueness lies in it's method of mitigating multipath interference. Alternating projection maximum likelihood is first used to estimate the AOA of the first received signal (the signal of interest) as well as that of all multipath arrivals (Kennedy et al., 1995). The receive antenna array is then electronically steered in such a way as to reject the multipath interference and provide maximum gain to the signal of interest. TDOA is estimated by cross-correlation of signals from two different cell sites and trilateration used to estimate the cellular telephone's position. The system is reportedly able to achieve a position accuracy of 100 m or better.

2.4 The CellocateÔ System

Cellocate™ is a TOA based cellular telephone positioning system currently implemented for AMPS. Using root MUSIC, the TOA of a signal transmitted by the cellular telephone is estimated by correlation at three cell sites or more. A GPS receiver, operating

in time-transfer mode, is located at each cell site to provide time synchronization. TOA estimates are differenced at a central processing site to produce TDOA measurements. A two dimensional position is then estimated by hyperbolic trilateration through the use of LS.

Since Cellocate™ makes use of operational cellular signals, it enjoys the advantages, and is subject to the limitations, of propagation in the 800 - 900 MHz portion of the radio spectrum. The system is capable of operation in all weather conditions and all propagation environments. However, as will be shown later, the performance of Cellocate™ is dependent on received SNR and multipath conditions.

2.4.1 The Cellular Registration Message

The signal used for correlation in AMPS is the preamble of the cellular telephone registration message. This message consists of a 30 bit dotting sequence, an 11 bit Barker coded synchronization word, and a 7 bit digital color code for a total of 48 bits. It is transmitted by each cellular telephone on a reverse control channel whenever network access is required. Each cell site makes use of two 30 kHz wide control channels. The forward control channel is for transmission from the network to the cellular telephone whereas transmission in the opposite direction is performed over the reverse control channel. When a cellular telephone is turned on, it scans all available forward control channels and locks onto the strongest one. Transmissions on the forward control channel are periodic. After locking onto a control channel, the cellular telephone receives data on that channel which among other things tells it when the reverse control channel for that particular cell is free. Upon availability of the reverse control channel, the cellular telephone transmits its registration message. Other occasions for registration are paging by the network, call initiation, and regular registration while powered up. A cellular telephone may be made to register as often as desired by paging it.

Because each cell has only one reverse control channel, all cellular telephones in that cell must share that channel. A positioning system must sort through all of the incoming transmissions on the reverse control channel and identify that of the cellular telephone of interest. This may be accomplished using the mobile identification number which is part of

the registration message.   More details regarding the structure and modulation of the registration message preamble are given in section 5.2.1.

2.4.2 RF Receiver and Correlator

In order to receive the registration message and estimate its TOA, a radio frequency (RF) correlation receiver and a GPS receiver are installed at each cell site. A block diagram of the system at each cell site is shown in Figure 2.4. The standard cell site equipment consists of a preselect filter, a low noise amplifier (LNA), and a splitter.
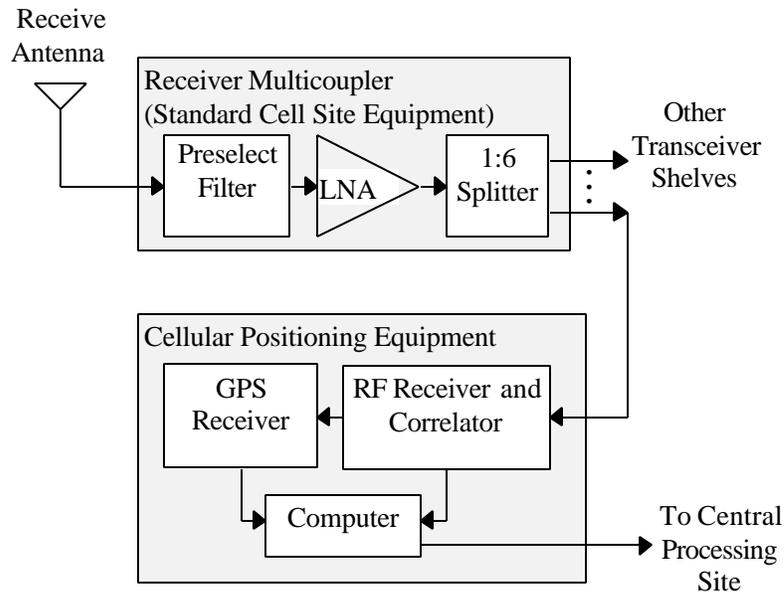
Figure 2.4   Cellocate™ System at a Cell Site

The Cellocate™ prototype consists of an RF receiver, a correlator, a GPS receiver, and a computer (Klukas et al., 1996). The RF receiver and correlator consists of an RF board and a digital signal processor (DSP). The system interfaces with the cellular network at the splitter in the receiver multicoupler as shown in Figure 2.4. The first task of the RF receiver, therefore, is to demodulate the signal. The modulation method used in AMPS is narrowband, analog FM. The signal is demodulated as shown in Figure 2.5.
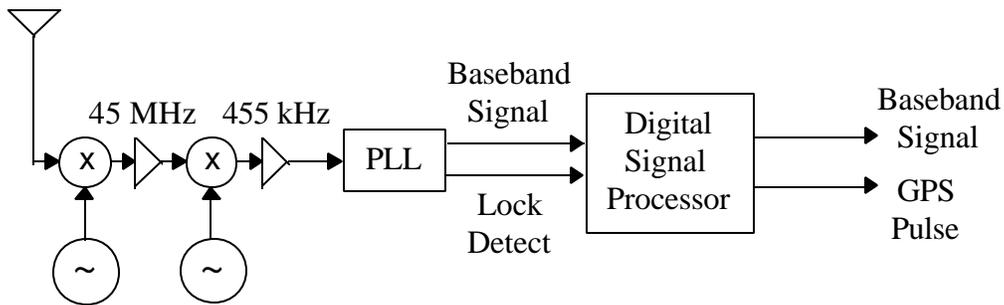
Figure 2.5　　　　RF Receiver and Correlator

The RF receiver is tuned to the appropriate reverse control channel. The signal is downconverted first to 45 MHz and then to 455 kHz. The FM signal is demodulated with a phase lock loop (PLL). The PLL provides two outputs, the baseband signal and the lock detect. The lock detect signal becomes active when the PLL has locked onto the received signal.

The baseband signal is sampled in the DSP at 160 kHz. The symbol rate of the registration preamble is 10 kbps. This gives 16 samples per data symbol for a total of 768 samples. When the lock detect signal becomes active, correlation on the Barker sync word is initiated. As each new sample is received, the contents of a correlation buffer are shifted over by one sample and correlated with a stored replica of the Barker word. Once a correlation value above a certain threshold has been detected, a counter counts up to 112 data samples, the known number of samples from the Barker correlation peak to the end of the preamble. If another correlation value greater than the threshold is detected before the counter reaches 112, the counter is reset.

When the counter reaches 112, a pulse is sent to the GPS receiver and the entire baseband signal, which has been stored in a separate buffer, is downloaded to the controlling cell site computer. When pulsed, the GPS receiver records a time stamp in GPS time to the same computer. Within the computer a second correlation is performed in software using the entire registration preamble and root MUSIC. When referenced to the GPS time stamp, the resulting correlation peak estimates the TOA of the signal at that cell site in GPS time. In this prototype system, MUSIC is not implemented in the DSP due to the complexity of the eigenanalysis required. The DSP in this prototype is not of sufficient speed to perform all of the required computations in the time allocated. In an operational system, it is possible for

registration messages from various cellular telephones to arrive within fractions of a second of each other. The system must be able to process each registration in turn. More powerful DPSs, which should be able to handle the processing load of MUSIC, are now becoming commercially available.

A second interface with the cellular network is required in order to transmit the TOA estimate, mobile identification number and cell site identification to a central site for position estimation. One possibility is to transmit this information over a land telephone line via modems.

2.4.3 GPS Time Synchronization

Crucial to any TOA based system is time synchronization. If TOA measurements from various locales are to be combined to estimate position, they must obviously be measured with respect to the same time reference. One possibility is to use high performance atomic clocks at the cell sites. However, compared to GPS time-transfer, atomic clocks have some serious disadvantages (Van Dierendonck et al., 1983). The first is cost. Since the construction of atomic clocks is labor intensive, the cost of these clocks is expected to remain high in the future. On the other hand, GPS receivers are continually dropping in cost due to the use of large scale integration techniques. Another disadvantage of frequency standards such as atomic clocks is their drift, without bound, over time. This necessitates periodic calibration. GPS receivers operating in time-transfer mode have no such problem and in the short term are just as accurate as a Cesium clock. One quickly concludes that GPS time-transfer is the most cost effective and accurate method of time synchronization for the positioning system under consideration.

GPS time-transfer is most accurate when the position of the GPS receiver is precisely known. Assuming that the receiver position is known and given that satellite positions are obtained from the navigation message in the GPS signal, the propagation distance between the receiver and the satellites may be calculated. The TOT of the GPS signal, in GPS system time, is also contained in the navigation message and thereby known by the receiver. This allows for calculation of the GPS signal TOA at the GPS receiver by adding the TOT to the known propagation distance. The TOA of the GPS signal in the

receiver's own time is then compared to the GPS system TOA just calculated to determine the receiver's time offset from true GPS time. In this way a GPS receiver operating in time-transfer mode at each cell site synchronizes all sites to GPS time.

In the Cellocate™ system, the position of each cell site is accurately surveyed beforehand in DGPS (Differential GPS) carrier phase mode with an accuracy better than 10 cm (Lachapelle et al., 1992). When operating in time-transfer mode, the receiver's coordinates are held fixed allowing all satellite observations to be used for solving the receiver's clock offset. The accuracy of the receiver's time, in the absence of SA and with known receiver coordinates, will be better than 30 ns (Van Dierendonck et al., 1983). The presence of SA increases the user equivalent range error. This corresponds to a timing error of $SA_{error}/\sqrt{n}$ where SA_{error} is the error in range due to SA and n is the number of satellites used. A nominal value for the standard positioning service timing error, in the presence of SA, is 280 ns 2 dRMS (probability of 95% - 98%) (National Research Council, 1995).

In the Cellocate™ system, coordinated time synchronization is possible (Van Dierendonck et al., 1983). The common mode - common view technique may be used to improve time-transfer accuracy analogous to the way DGPS improves positioning accuracy. If the GPS receivers at all cell sites are constrained to view the same satellites, common errors will be reduced or even eliminated over short baselines when TOA differences are formed in the hyperbolic trilateration process. Common errors include satellite ephemeris and clock errors (including SA) as well as atmospheric errors. Errors not common to all receivers will of course be amplified by $\sqrt{2}$. This includes the above mentioned satellite and atmospheric errors for long baselines where these errors are weakly correlated between receivers. For short baselines ($\leq$ 100 km), the error budget will be dominated by receiver noise and error in the receiver coordinates. As a result, for short baselines and cm level relative receiver coordinate accuracy, a relative timing accuracy of $\leq$ 5 ns between cell sites may be possible.

A simple test was performed to see how closely two GPS receivers could be synchronized in time. Both receivers used were the NovAtel GPSCard™, a C/A code receiver incorporating Narrow Correlator™ spacing (Fenton et al., 1991). NovAtel GPSCard™ receivers were used for all GPS work in this project. The two receivers shared the same GPS antenna. This is referred to as a zero baseline test and is extremely useful for isolating GPS receiver errors. Both receivers were also pulsed by the same source and when pulsed recorded a time stamp to a file. A total of 657 time stamps were obtained over approximately 32 minutes. The differences between corresponding time stamps are plotted in Figure 2.6.
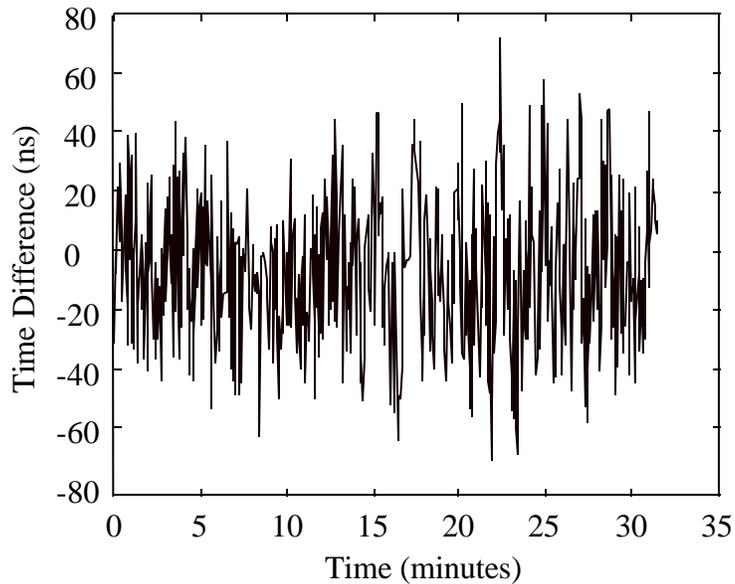
Figure 2.6   Time Stamp Differences between Two GPS Receivers

The mean time difference in Figure 2.3 is -6.5 ns and the standard deviation is 23.8 ns. This is significantly higher than the 5 ns just quoted. This is due to the resolution of the time stamp measurement in the GPSCard™. According to the Commands Description Manual (NovAtel, 1994) the resolution is 49 ns. This suggest a maximum error of half that, or 24.5 ns. Although quantization error is uniformly distributed, adding two uniformly

distributed random variables results in a nearly Gaussian random variable (Maybeck, 1994).

The standard deviation of the time differences is then expected to be $0.67*24.5*\sqrt{2} = 23.2$ ns. Time synchronization may therefore be expected to contribute approximately 7 m (67%) of error to the TDOA measurements using a GPS receiver of the NovAtel GPSCard™ type.

2.4.4 Position Estimation

Cellocate™ employs hyperbolic trilateration as discussed in section 2.2. A minimum of two TDOAs are required in the horizontal case since there are two unknowns, latitude and longitude. In the case where the number of observations (TDOAs) exceeds the number of unknowns, errors in the observations prevent a unique solution. LS is a common and effective method of dealing with this redundancy by minimizing the sum of the square of the residuals (Krakiwsky, 1990). The hyperbolic trilateration equations are nonlinear and must be linearized for use in LS. The linearization process may be implemented with a Taylor series expansion. Successive iterations of LS are then used to arrive at a position estimate. Other methods exist which linearize a non-linear problem without the need for iterations. Such closed form solutions include the spherical interpolation, spherical intersection, and plane intersection methods (Smith et al., 1987) as well as the feasible bivector method (Schmidt, 1996). Plane intersection is herein investigated for providing the initial position to begin the iterative process of LS. Both plane intersection and LS are described in detail in Chapter 4.

CHAPTER 3

TOA ESTIMATION

3.1 Introduction

In Chapter 2, radio location by TOA estimation was presented as a leading candidate to solve the cellular telephone positioning problem. Correlation of signals was discussed as a common method of estimating TOA. This chapter deals in particular with TOA estimation by correlation. It begins by showing how TOA estimation is actually the estimation of the channel impulse response and how correlation is related to the input/output relations of linear systems. This is followed by a discussion of MUSIC and its application to the correlation process.

3.2 The Channel Impulse Response

In a cellular telephone communication system, signals are transmitted through a channel and thereby corrupted by noise and very often, multipath. Consider the simple block diagram of Figure 3.1.
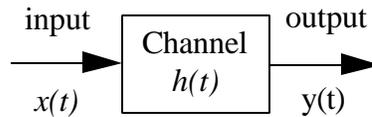
Figure 3.1   Cellular Communication Channel

In Figure 3.1, the signal transmitted by the cellular telephone is *x(t)*. It passes through the channel with impulse response *h(t)* and the signal arriving at the cell site is *y(t)*. The impulse response of a system is defined as the output of the system when the input is a unit impulse (Van Valkenburg, 1964). The Dirac Delta function, $\delta(t)$, is commonly used to represent a unit impulse and is illustrated in Figure 3.2. If *x(t)* = $\delta(t)$, then *y(t)* will equal the impulse response, *h(t)*. The channel is characterized in the frequency domain by its

frequency response, *H(f)*. The impulse response *h(t)* and frequency response *H(f)* are a Fourier transform pair.

$$\delta(t - t_d)$$
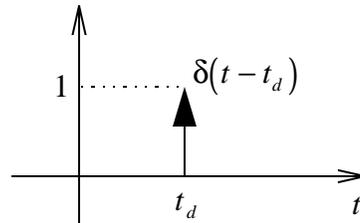
Figure 3.2   The Dirac Delta Function

The output of the channel is related to the input by the impulse response. This is expressed in the convolution integral (Proakis, 1989)

$$y(t) = \int_{-\infty}^{\infty} h(\tau) x(t - \tau) d\tau.$$

(3.1)

Rather than solve the convolution integral, it is often more convenient to make use of the convolution theorem (Haykin, 1989). This theorem states that convolution in the time domain and multiplication in the frequency domain are a Fourier transform pair. Rather than convolve two time domain signals, one may multiply their individual Fourier transforms in the frequency domain. Inverse transforming back to the time domain then gives the same result as (3.1).

3.3 Impulse Response of Multipath Channels

In TOA estimation we are not interested in determining the channel output *y(t)* but the channel impulse response *h(t)*. Due to multipath, the urban radio propagation channel

for cellular frequencies is generally modeled by a linear filter (Turin et al., 1972a). If *x(t)* is the transmitted lowpass waveform, then the received lowpass signal, $y(t)$, is

$$y(t) = \sum_{k=0}^{N-1} a_k x(t - t_k) e^{j\theta_k} + n(t)$$

(3.2)

where N = the number of multipath arrivals,

a_k = the amplitude of arrival k,

t_k = the time delay of arrival k,

θ_k = the phase of arrival k, and

n(t) = additive Gaussian noise.

Of interest here is that the received signal is the sum of a number of scaled, phase shifted, and time delayed replicas of the transmitted signal. If *x(t)* = $\delta(t)$, then $\left| y(t) \right|$, the magnitude of *y(t)*, and hence the magnitude of the impulse response *h(t)*, may appear something like Figure 3.3.
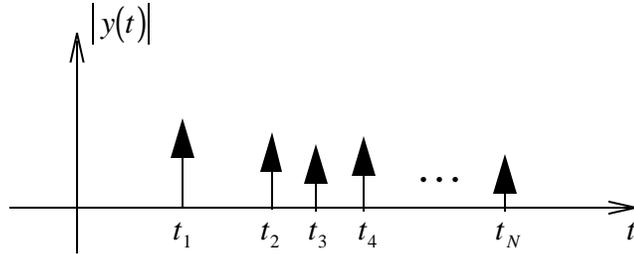
Figure 3.3   Typical Impulse Response Magnitude of Multipath Channel

For radio location purposes, the TOA of the first, and hopefully LOS arrival, t_1, is the quantity of interest. TOA estimation may then be thought of as estimating the impulse response of the channel, identifying the first arrival and its time delay. If the individual arrivals in the impulse response are resolved and the first arrival corresponds to the LOS path, multipath introduces no error into the TOA estimate. Many times, however, the

25

individual arrivals are so close that they cannot be resolved. In that case, the first signal in the estimated impulse response actually consists of a number of paths and its time delay will be skewed due to the later arrivals. Therefore, high resolution in the TOA estimation process is an important factor.

The impulse response of the channel may be considered to consist of a number of point sources in white noise. As will be discussed shortly, this is an important condition as far as MUSIC is concerned.

3.4 Estimating the Channel Impulse Response

To estimate the channel impulse response *h(t)*, we may cross-correlate the input of the channel with the output. This assumes that the input to the channel, the transmitted signal, is known. Following Proakis (1989), the cross-correlation function between the input and output of a linear system is found. For jointly stationary stochastic processes *x(t)* and *y(t)*, their cross-correlation is,

$$\phi_{yx}(t_1, t_2) = E(Y_{t_1} \cdot X_{t_2}) = \int_{-\infty}^{\infty} h(\alpha) E\big[X(t_1 - \alpha) X(t_2) \big] d\alpha$$

$$= \int_{-\infty}^{\infty} h(\alpha)\phi_{xx}(t_1 - t_2 - \alpha)d\alpha$$

$$= \phi_{yx}(t_1 - t_2).$$

(3.3)

With $t_1 - t_2 = \tau$, the cross-correlation becomes

$$\phi_{yx}(\tau) = \int_{-\infty}^{\infty} h(\alpha)\phi_{xx}(\tau - \alpha)d\alpha \qquad (3.4)$$

which is a convolution. We may apply the convolution theorem to express (3.4) in the frequency domain. The result is

$$\Phi_{yx}(f) = \Phi_{xx}(f)H(f) \qquad (3.5)$$

where $\Phi_{yx}(f)$ is the Fourier transform of $\phi_{yx}(t)$ and $\Phi_{xx}(f)$ is the energy density spectrum of the input signal $x(t)$. The energy density spectrum of a signal is related to its autocorrelation function by

$$\Phi_{xx}(f) = \Im\left[\phi_{xx}(t)\right] = \Im\left[x(t)* x^*(-t)\right] \qquad (3.6)$$

where $\Im$ represents the Fourier transform, $*$ indicates convolution, and $x^*(t)$ is the complex conjugate of $x(t)$. If the input signal is known, its autocorrelation function, and hence energy density spectrum, may be determined.

It then remains to determine the cross-correlation between the input and output signals without knowledge of the channel transfer function. The cross-correlation between two signals may also be found from

$$\phi_{xy}(\tau) = \int_{-\infty}^{\infty} x(t) y^*(t - \tau) dt$$

$$= x(\tau)* y^*(-\tau). \qquad (3.7)$$

Again it is generally easier to implement multiplication in the frequency domain than convolution in the time domain. Hence, making use of the convolution theorem we have

$$\phi_{xy}(\tau) = \Im^{-1}\left[\Phi_{xy}(f)\right] = \Im^{-1}\left[X(f)\cdot Y^*(f)\right] \quad (3.8)$$

where $\Im^{-1}$ is the inverse Fourier transform.

From (3.5) and (3.8) the impulse response of the channel, and hence TOA, is then

calculated by
$$h(t) = \Im^{-1}\left[\frac{X(f)\cdot Y^*(f)}{\Phi_{xx}(f)}\right]. \qquad (3.9)$$

This manner of estimating TOA is often referred to as the correlation method. Note that if the input signal $x(t)$ is white noise, $\Phi_{xx}(f)$ is a constant. Estimating TOA then reduces to finding the cross-correlation between the transmitted signal and the received signal.

3.5 TOA Estimation using Spread Spectrum Codes

Spread spectrum signals are natural candidates for estimating TOA by correlation. Spread spectrum signals are so named due to their wide bandwidth (Proakis, 1989). Pseudo-noise (PN) codes are impressed upon the signal to be transmitted spreading its bandwidth. The signal then appears to be pseudo-random or noise-like. The higher the chip rate of the PN code, the wider the bandwidth of the spread spectrum signal and the more it appears as noise.

The noise-like character of spread spectrum signals may be seen by observing the autocorrelation function of PN codes. The most common type of codes used for spread spectrum signals are maximum-length shift-register sequences (Proakis, 1989). For a code of length n, the autocorrelation function will appear as in Figure 3.4. T_c is the chip length and the base of the autocorrelation peak is twice T_c. As the chip length decreases, the autocorrelation function begins to resemble an impulse and the code appears to be more and more noise-like.   As discussed in Chapter 2, a higher chip rate, and hence a larger bandwidth, improves TOA estimation in the presence of multipath.
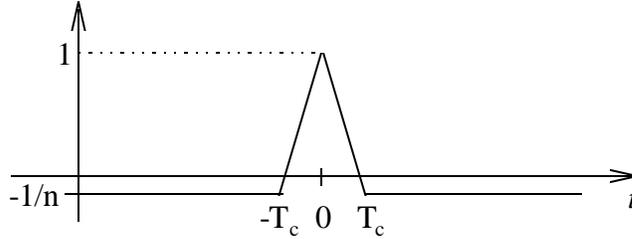
Figure 3.4   Autocorrelation Function of PN Code

For PN codes with large periods, the energy density spectrum will approximate the sinc^2 function. The energy density spectrum of the autocorrelation function depicted in Figure 3.4 is

$$\Phi_{xx}(f) = T_c \left[\frac{\sin(T_c \pi f)}{T_c \pi f} \right]^2. \qquad\qquad (3.10)$$

Equation (3.10) is plotted in Figure 3.5. Note that as the chip rate increases (T_c decreases), the nulls in (3.10) move away from zero and in the extreme, $\Phi_{xx}(f)$ becomes flat (white noise).
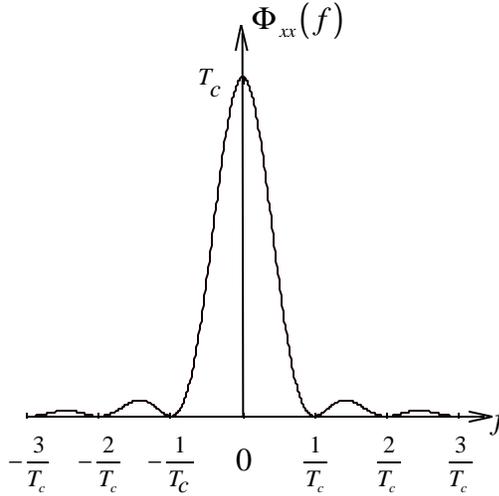
Figure 3.5　　　　　Energy Density Spectrum of PN Code

For the purpose of estimating the TOA of the cellular registration preamble, the actual data transmitted acts as the PN code. However, the data rate is only 10 kbps. Therefore, $\Phi_{xx}(f)$ in (3.9) is not a constant and must be accounted for if the true channel impulse response is to be estimated. Dividing by $\Phi_{xx}(f)$ removes the autocorrelation spectrum from the cross-correlation. The result, in the time domain, are impulses at times corresponding to the TOA of various arrivals. This process is often called inverse filtering or deconvolution and may also be used to remove the effects of filtering. Note that prior to dividing by $\Phi_{xx}(f)$, the signals involved must be bandlimited in order to avoid dividing by the nulls in Figure 3.5. Noise is enhanced when dividing by the nulls.

At this point it is possible to demonstrate, in the frequency domain, the effect of multipath on TOA estimation. In the absence of multipath, the ideal channel impulse

response will consist of only one impulse. The frequency response will, as a consequence, be flat. For a two ray channel the frequency response will appear something like Figure 3.6 (Dumont, 1994). The nulls in the magnitude of the frequency response are due to destructive interference between the two arrivals and cause fading of the resultant signal. The distance between the nulls is the inverse of the time delay between the two arrivals. The location of the nulls is a function of their phases.
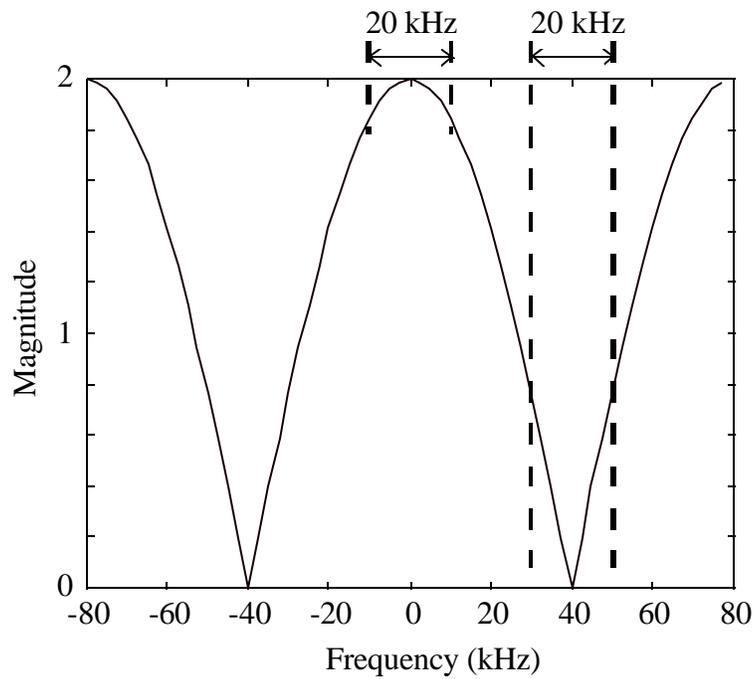
Figure 3.6   Frequency Response of Two Ray Channel

As discussed above, in the process of estimating the channel impulse response, the signals involved are bandlimited. The vertical dotted lines in Figure 3.6 indicate, for a data rate of 10 kHz and a sampling frequency of 160 kHz, those portions of the spectrum left after bandlimiting. The 20 kHz portion of spectrum used for TOA estimation is always centered at 0 Hz. If the signal phases are such that the frequency spectrum appears as in Figure 3.6, it

may be difficult to resolve the two arrivals since the frequency response is somewhat flat. As the two arrivals come closer together in time, the nulls in Figure 3.6 move further apart resulting in an even flatter 20 kHz spectrum. The result is a spectrum which begins to resemble that of a one ray channel. On the other hand, if the 20 kHz spectrum centered at 0 Hz is also centered on a fade, a poor SNR will result in only noise being used for the TOA estimation process. This will obviously give poor results. It is, therefore, desirable that the 20 kHz portion of spectrum used for TOA estimation be neither centered on a peak nor a fade. Unfortunately, this cannot be guaranteed since the location of the nulls with respect to 0 Hz is a function of phase and phase is a uniformly distributed random variable.

3.6 The Correlation Method

A block diagram of the correlation method is presented in Figure 3.7. The frequency domain signals are multiplied and the result is multiplied by a weighting function $W(f)$. To this point, the weighting function discussed has been $W(f) = 1/\Phi_{xx}(f)$. In the literature (Carter, 1993), TOA estimation using this weighting function is termed the Roth method after (Roth, 1971).
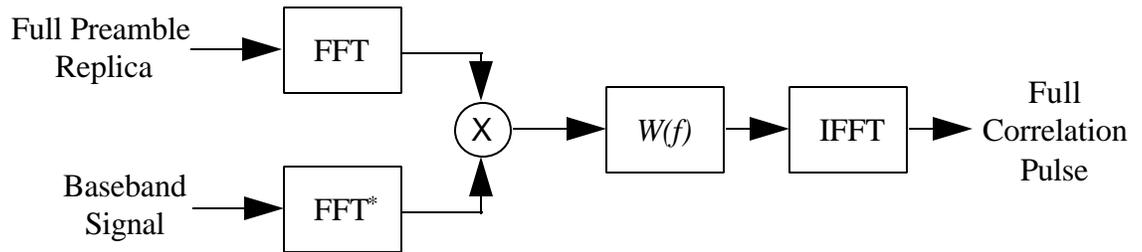
Figure 3.7   Correlation Method

Other weighting functions have been derived and investigated as well (Carter, 1993). They are summarized in Table 3.1. The parameter $C_{xy}(f)$ is called the magnitude-squared coherence and is defined as

$$C_{xy}(f) = \frac{\left|\Phi_{xy}(f)\right|^2}{\Phi_{xx}(f)\Phi_{yy}(f)}. \qquad\qquad (3.11)$$

All of the weighting functions in Table 3.1 are merely methods of manipulating the frequency spectrum in order to minimize the effects of noise. They do not overcome the fundamental limitation of the discrete Fourier transform - the inability to resolve the TOA of a signal lying between samples. Both standard cross-correlation and the Roth method were tested with simulations and the results were identical to within a millimetre.

Table 3.1   Weighting Functions for the Correlation Method

Method	$W(f)$
Standard Cross-Correlation	1
Roth	$1/\Phi_{xx}(f)$
Wiener Processor	$C_{xy}(f)$
Smoothed Coherence Transform	$1/\sqrt{\Phi_{xx}(f)\Phi_{yy}(f)}$
Phase Transform	$1/\left|\Phi_{xy}(f)\right|$
Maximum Likelihood	$\dfrac{C_{xy}(f)}{\left(1-C_{xy}(f)\right)\left|\Phi_{xy}(f)\right|}$

3.7 Correlation in the Cellocate Ô System

3.7.1 Full Correlation

In section 2.4.2, acquisition and time tagging of the cellular registration preamble by the Cellocate™ system was discussed. Sliding correlation on the Barker word is used to give an initial TOA estimate in GPS time. The entire baseband signal is then downloaded to the site computer for further processing. This processing includes TOA estimation by the

correlation method discussed in sections 3.4, 3.5, and 3.6. The peak of the full correlation pulse of Figure 3.7 may be referenced to the GPS time tag previously obtained. The result is a Fourier based TOA estimate. For the application at hand, position estimation based on such TOA estimates is not sufficiently accurate, particularly in the presence of multipath. This is demonstrated in Chapter 8 where the standard cross-correlation method of Table 3.1 is used on some of the data. Superresolution algorithms such as MUSIC may be used at this point to improve the TOA estimation accuracy.

3.7.2 Correlation with root MUSIC

3.7.2.1 Standard MUSIC

MUSIC is widely discussed in the literature. A concise summary is given here for the convenience of the reader. The development below basically follows that of Haykin (1991), Klukas (1993), and Dumont (1994).

MUSIC is an eigenvector based superresolution algorithm initially intended for frequency and direction of arrival estimation (Schmidt, 1986). MUSIC models frequency domain data as point sources in noise and therefore assumes that the time domain data consists of complex sinusoids in noise. For frequency estimation, the time domain data is processed by means of a transversal filter. Hence, the data on which MUSIC operates is an equally spaced time series.

Consider a signal consisting of L uncorrelated sinusoids in additive white noise with zero mean and variance σ^2. Let $\mathbf{R}$ be the $(M+1)$ by $(M+1)$ ensemble averaged correlation matrix of the signal. It may be expressed in the form

$$\mathbf{R} = \mathbf{SDS}^H + \sigma^2 \mathbf{I} \qquad (3.12)$$

where H indicates Hermitian transpose and $\mathbf{I}$ is the $(M+1)$ by $(M+1)$ identity matrix. $\mathbf{S}$ is an $(M+1)$ by L matrix of the form

$$S = [s_1, s_2, \cdots, s_L]$$

$$= \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{-j\omega_1} & e^{-j\omega_2} & \cdots & e^{-j\omega_L} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-jM\omega_1} & e^{-jM\omega_2} & \cdots & e^{-jM\omega_L} \end{bmatrix} \tag{3.13}$$

and is called the frequency matrix. $\mathbf{D}$ is the correlation matrix of the sinusoids and therefore of dimension L by L. Since the sinusoids are assumed to be uncorrelated, $\mathbf{D}$ is diagonal, nonsingular, and of rank L. Let the eigenvalues of $\mathbf{SDS}^H$ be $\{\upsilon_1 \geq \upsilon_2 \geq \cdots \geq \upsilon_{M+1}\}$ and those of $\mathbf{R}$ be $\{\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{M+1}\}$. Then according to (3.12)

$$\lambda_i = \upsilon_i + \sigma^2, \quad i = 1, 2, \cdots, M+1. \tag{3.14}$$

If the L sinusoids have distinct frequencies, $\mathbf{S}$ will also be of rank L. As a result $\mathbf{SDS}^H$ will be of rank L and have L nonzero eigenvalues. Therefore, the eigenvalues of $\mathbf{R}$ will be

$$\lambda_i = \begin{cases} \upsilon_i + \sigma^2, & i = 1, \cdots, L \\ \sigma^2, & i = L+1, \cdots, M+1. \end{cases} \tag{3.15}$$

According to (3.15), the eigenvalues may be partitioned into two sets, those that correspond to the signals and those that correspond to noise. The corresponding eigenvectors may be similarly partitioned. Let $\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{M+1}\}$ be the eigenvectors of $\mathbf{R}$. Due to the fact that

$$\left(\mathbf{R} - \sigma^2 \mathbf{I}\right)\mathbf{v}_i = \mathbf{0}, \quad i = L+1, \, \cdots, \, M+1 \qquad (3.16)$$

the columns of $\mathbf{S}$ are related to the eigenvectors corresponding to the $M+1$ - L smallest eigenvalues of $\mathbf{R}$ by (Haykin, 1991),

$$\mathbf{s}_l{}^H \mathbf{v}_i = 0, \quad i = L+1, \, \cdots, \, M+1$$
$$l = 1, \, 2, \, \cdots, \, L. \qquad (3.17)$$

MUSIC exploits (3.17) for the purpose of determining the frequencies of the signals. The observation space may be said to consist of two subspaces which are orthogonal to one another; the signal plus noise subspace and the noise subspace. Those eigenvectors corresponding to the $M+1$ - L smallest eigenvalues span the noise subspace. The columns of $\mathbf{S}$ span the signal plus noise subspace and according to (3.17) are orthogonal to the noise subspace. Let $\mathbf{s}(\omega)$ be a variable frequency scanning vector with form

$$\mathbf{s}^T(\omega) = \left[1, e^{-j\omega}, \cdots, e^{-jM\omega}\right]. \qquad (3.18)$$

Any $\mathbf{s}(\omega)$, corresponding to the frequency of one of the L signals, when projected onto the noise subspace, will result in a null. Frequency estimation by MUSIC may then be achieved by scanning (3.18) through all possible frequencies and projecting onto the noise subspace. The MUSIC spectrum so obtained is calculated from

$$\hat{S}_{\mathrm{MUSIC}}(\omega) = \frac{1}{\mathbf{s}^H(\omega)\mathbf{V}_N \mathbf{V}_N{}^H \mathbf{s}(\omega)} \qquad (3.19)$$

where $\mathbf{V}_N = \left[\mathbf{v}_{L+1}, \cdots, \mathbf{v}_{M+1}\right].$

3.7.2.2 Decorrelation of Signals

MUSIC relies on the assumption that the signals to be estimated are uncorrelated. If some of the sinusoids represented by (3.13) are partially correlated, **D** is no longer diagonal but still nonsingular and of rank L. In the event that some of the signals are perfectly correlated (coherent), **D** becomes singular and rank deficient. When this happens **R** is no longer of rank L and (3.15) and (3.17) no longer hold.

Practical experience suggests that the signals need not be perfectly correlated to cause problems (Shan et al., 1985). Highly correlated signals created by a multipath environment also have detrimental effects on the performance of MUSIC. To use MUSIC in an application such as cellular positioning where multipath may be considered to be always present, requires decorrelation of the signals.

Signal decorrelation by smoothing of the data has been found to be very effective (Pillai et al., 1989). Consider the N data points $\left[x(0)\, x(1)\, \cdots\, x(N-1) \right]$. A subarray is formed consisting of $M+1$ points where $M+1 < N$. Smoothing is introduced by sliding the subarray across the full array of data points in both the forward and backward directions. This is accomplished by structuring a data matrix **A** in the form,

$$
\mathbf{A}^{H} = \left[
\begin{array}{ccc|ccc}
x(M) & \cdots & x(N-1) & x^{*}(0) & \cdots & x^{*}(N-M+1) \\
x(M-1) & \cdots & x(N-2) & x^{*}(1) & \cdots & x^{*}(N-M+2) \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
x(0) & \cdots & x(N-M+1) & x^{*}(M) & \cdots & x^{*}(N-1)
\end{array}
\right].
$$

(3.20)

The left half of $\mathbf{A}^{H}$ contains the subarrays for forward smoothing whereas the right half consists of the subarrays for backwards smoothing.

The ensemble averaged correlation matrix **R** is defined as the expected value of the product of the data matrix **A**, with its Hermitian. In practice it is not known and must be estimated from a sample average. The estimate of **R** will then be

$$\hat{\mathbf{R}} = \frac{1}{2(N-M)} \mathbf{A}^H \mathbf{A} \qquad (3.21)$$

where $2(N-M)$ is the number of points averaged.

3.7.2.3 root MUSIC

To avoid the frequency scanning necessary to obtain the spectrum of (3.19), root MUSIC was developed (Haykin, 1991). In this approach, a polynomial $D(z)$ is formed from the denominator of (3.19) by substituting $e^{j\omega}$ in (3.18) with the complex variable z. The coefficients of the polynomial are equal to the sums of the diagonals of the matrix $\mathbf{V}_N \mathbf{V}_N{}^H$. Those roots of $D(z)$ which are closest to the unit circle should correspond to signals and the phase of those roots to the normalized frequencies of those signals.

Simulation results by (Rao et al., 1989) demonstrate that root MUSIC is preferable to standard MUSIC due to the radial nature of the errors in root MUSIC. A radial error in root location does not effect the phase of the root as it does the corresponding peak in the standard MUSIC spectrum. The result is a more accurate frequency estimate with root MUSIC.

3.7.2.4 Time Estimation with MUSIC

Although MUSIC was originally applied to the problem of frequency estimation, it is equally suitable for time estimation. Yamada et al. (1991) are one of the first to use MUSIC in this way. Consider the sampled time domain signal

$$x(i) = \sum_{l=1}^{L} \alpha_l e^{j\omega_l i} + v(i), \quad i = 0,\, 1,\ldots, N-1 \qquad (3.22)$$

where L = the number of complex sinusoids,

α_l = the complex amplitude of the lth sinusoid,

ω_l = the angular frequency of the lth sinusoid, and

$v(i)$ = the complex noise value at the ith sampling time.

In (3.22) the angular frequencies ω_l, are the parameters to be estimated and the data points $x(i)$ differ from one another in terms of time. The data represented by (3.22) is used to form the data matrix of (3.20) and the angular frequencies are estimated as described above.

Now consider the problem of estimating time delays from frequency domain data collected by a network analyzer. To measure the frequency response of a system, a network analyzer injects a sinusoid into the system and sweeps the frequency of the sinusoid through the band of interest. Reflections of the sinusoid from the system at each frequency are the measured data of interest here. Following Yamada et al. (1991) the signal at any frequency $f(i)$ may be written as

$$r(i) = \sum_{k=1}^{L} s_k \cdot e^{-j2\pi f(i)t_k} + n(i), \quad i = 0,\ 1,\ldots,\ N-1 \qquad (3.23)$$

where L = the number of reflections,

s_k = the reflection coefficient of the kth reflection,

t_k = the time delay of the kth reflection, and

$n(i)$ = additive white noise at frequency $f(i)$.

Because (3.23) represents frequency domain data, the time delays t_k are the parameters to be estimated and the data points are a frequency series indexed by $f(i)$. However, given that $\omega = 2\pi f$, note the similarity between (3.22) and (3.23). If $f(i)$ in (3.23) is replaced with the index i, then (3.22) and (3.23) have the same form with the exception of the domain in which the signal is represented. To estimate the time delays of

(3.23), a time scanning vector, analogous to the frequency scanning vector (3.18), is used. It takes the form of

$$\mathbf{s}^T(t) = \left[1, e^{-jt}, \cdots, e^{-jMt} \right] \qquad (3.24)$$

where M is the number of frequency points used in a subarray of data. The data matrix (3.20) and correlation matrix estimate (3.21) may then be formed from the N frequency domain points. Note that decorrelation by smoothing is necessary due to the fact that the reflections are highly correlated. Following eigen-decomposition of the correlation matrix, the MUSIC spectrum is calculated from (3.19) with $\mathbf{s}(t)$ in place of $\mathbf{s}(\omega)$. The root MUSIC algorithm may also be used in the frequency domain merely by substituting z for e^{jt} in (3.24).

We see then that MUSIC is able to estimate time from frequency domain data just as it is able to estimate frequency from time domain data. In both cases, the parameters to be estimated are modeled as point sources in noise. For frequency estimation this requires the signal to consist of sinusoids and for time estimation the signal must consist of a number of impulses. Therefore, we return to the interpretation that TOA estimation is, in reality, the estimation of the channel impulse response.

3.7.2.5 TOA Estimation in Cellocate Ô

Section 3.7.1 described how a Fourier based TOA estimate is first generated by the Cellocate™ system through correlation of the entire registration message preamble. This process is identical to that illustrated in Figure 3.7 with $W(f) = 1$. This corresponds to the standard cross-correlation method. In order to improve the resolution and accuracy of the correlation peak, the data is further processed with root MUSIC as in Figure 3.8.

Full Correlation Pulse → FFT → W(f) → root MUSIC → High Resolution Correlation Pulse
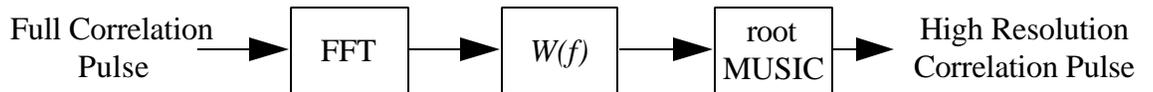
Figure 3.8   High Resolution Correlation by root MUSIC

The full correlation pulse used as input to Figure 3.8 was first windowed around the correlation peak. This was necessary to ensure accurate representation in the frequency domain following the FFT. Following transformation into the frequency domain by the FFT, the data is filtered by

$$W(f) = \frac{1}{\Phi_{xx}(f)H_F(f)} \qquad (3.25)$$

where $H_F(f)$ is the combined frequency response of any filters.

As discussed in section 3.5, this filtering removes the autocorrelation spectrum as well as any filter effects from the cross-correlation. What remains is the bandlimited impulse response of the channel. Since the impulse response consists of point sources (impulses), the data model is that assumed by MUSIC. The frequency domain data may then be used to form a data matrix and root MUSIC implemented as discussed previously. Details regarding the implementation of root MUSIC for TOA estimation in the Cellocate™ system are covered in section 5.4.

CHAPTER 4

POSITION ESTIMATION

4.1 Introduction

In section 2.4.4, LS was introduced as a method of implementing trilateration in the presence of observation redundancy. LS is used to make redundant observations consistent while satisfying a minimum variance criterion. This chapter explains the models used for trilateration, their linearization, and the mechanism of LS for estimating position. This is followed by a discussion of reliability and statistical testing of the LS results. The chapter then concludes with an explanation of plane intersection, a closed form position estimation algorithm.

4.2 Least Squares

4.2.1 The General Case

According to Krakiwsky (1990), "Least squares estimation is the standard method to obtain a unique set of values for a set of unknown parameters ($\mathbf{x}$) from a redundant set of observables ($\mathbf{l}$) through a known mathematical model ($\mathbf{f(x,l)}$)". Following Krakiwsky (1990), we consider the most general situation in which the parameters and observables are related through an implicit and nonlinear mathematical model. In that case we may write,

$$\mathbf{f}\left(\hat{\mathbf{x}},\hat{\mathbf{l}}\right) = 0 \qquad (4.1)$$

where $\hat{\mathbf{x}}$ and $\hat{\mathbf{l}}$ are estimates of the parameters and observations, respectively. These estimates are obtained with the condition that

$$\left[\hat{\mathbf{r}}^T\mathbf{C}_\mathbf{l}^{-1}\hat{\mathbf{r}} + \hat{\mathbf{d}}^T\mathbf{C}_\mathbf{x}^{-1}\hat{\mathbf{d}}\right] = \text{minimum}, \qquad (4.2)$$

where $\hat{\mathbf{r}}$ = estimated residuals to the observations,

$\hat{\mathbf{d}}$ = estimated corrections to the parameters,

$\mathbf{C_l}$ = the covariance matrix of the observations, and

$\mathbf{C_x}$ = the covariance matrix of the parameters.

LS does not require that the observations be of any particular distribution or that the distribution be known a priori (Mikhail et al., 1976). However, the observations should be zero mean. All systematic errors should be accounted for in the mathematical model. Also, for the condition of (4.2) to ensure that the variances of the estimated parameters are minimized, the observation errors must be symmetrically distributed. If the observation errors are normally distributed, the LS estimates will be identical to the maximum likelihood estimates.

The fact that corrections to the parameters are estimated instead of the parameters themselves, stems from the need to linearize the nonlinear model $\mathbf{f(x,l)}$. To do this, the mathematical model is approximated with a Taylor series (Swokowski, 1983). Discarding all second order and higher terms results in

$$\begin{aligned}\mathbf{f}\left(\hat{\mathbf{x}}, \hat{\mathbf{l}}\right) &= \mathbf{f}\left(\mathbf{x,l}\right) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x,l}} \left(\hat{\mathbf{x}} - \mathbf{x}\right) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\mathbf{x,l}} \left(\hat{\mathbf{l}} - \mathbf{l}\right) \\ &= \mathbf{f}\left(\mathbf{x,l}\right) + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x,l}} \hat{\mathbf{d}} + \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\mathbf{x,l}} \hat{\mathbf{r}} \;\; = \;\; 0.\end{aligned} \tag{4.3}$$

Note that $\mathbf{x}$ refers to the quasi-observed values of the parameters whereas $\hat{\mathbf{x}}$ refers to the adjusted or corrected parameters. In like manner, $\mathbf{l}$ refers to the observed values of the observables and $\hat{\mathbf{l}}$ to the adjusted observables. The value of the mathematical model for the observed and quasi-observed values of observables and parameters respectively, is called the misclosure vector, $\mathbf{w}$. We may then write,

42

$$\mathbf{w} = \mathbf{f}(\mathbf{x}, \mathbf{l}) \neq 0 \qquad (4.4)$$

which illustrates that given a set of redundant observations containing errors, a unique solution to the parameters is not possible. To arrive at a unique solution, the observations must be adjusted under the condition (4.2). In any case, the linearized mathematical model is expressed by

$$\mathbf{A}\hat{\mathbf{d}} + \mathbf{B}\hat{\mathbf{r}} + \mathbf{w} = \mathbf{0} \qquad (4.5)$$

where $\quad \mathbf{A} = \left.\dfrac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x},\mathbf{l}}\quad$ is the first design matrix and

$$\mathbf{B} = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{l}}\right|_{\mathbf{x},\mathbf{l}}\quad \text{is the second design matrix.}$$

The estimated corrections to the parameters and observations are determined from the LS normal equations. The normal equations themselves are derived from the linearized mathematical model and the variation function

$$\phi = \hat{\mathbf{r}}^T \mathbf{C_l}^{-1} \hat{\mathbf{r}} + \hat{\mathbf{d}}^T \mathbf{C_x}^{-1} \hat{\mathbf{d}} + 2\hat{\mathbf{k}}^T \left(\mathbf{A}\hat{\mathbf{d}} + \mathbf{B}\hat{\mathbf{r}} + \mathbf{w}\right) \quad (4.6)$$

where $\mathbf{k}$ is a vector of Lagrange correlates (Krakiwsky, 1990). Solving the LS normal equations for the parameter corrections results in

$$\hat{\mathbf{d}} = -\left[\mathbf{A}^T \left(\mathbf{B}\mathbf{C_l}\mathbf{B}^T\right)^{-1}\mathbf{A} + \mathbf{C_x}^{-1}\right]^{-1}\mathbf{A}^T \left(\mathbf{B}\mathbf{C_l}\mathbf{B}^T\right)^{-1}\mathbf{w} \;.(4.7)$$

The corrections or residuals to the observations are

43

$$\hat{\mathbf{r}} = -\mathbf{C_l}\mathbf{B}^T\hat{\mathbf{k}} \tag{4.8}$$

where $\quad \hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{C_l}\mathbf{B}^T\right)^{-1}\left(\mathbf{A}\hat{\mathbf{d}} + \mathbf{w}\right)$.

4.2.2 The Parametric Case

        The equations given above apply to the general case in which the relationship between the parameters and observations is implicit. For the case of trilateration, the observations may be explicitly expressed as a function of the parameters. This is known as the parametric case. Because the mathematical model is explicit in the observations, the second design matrix $\mathbf{B}$ is equal to $-\mathbf{I}$, the negative identity matrix. Given this as well as $\mathbf{C_x}^{-1} = 0$, meaning the parameters are unweighted, the LS equations are

$$\hat{\mathbf{d}} = -\left[\mathbf{A}^T\mathbf{C_l}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^T\mathbf{C_l}^{-1}\mathbf{w}^0 \tag{4.9}$$

where $\quad \mathbf{w}^0$ is the misclosure vector evaluated at some approximate

values of the parameters $\mathbf{x}^0$,

$$\hat{\mathbf{x}} = \mathbf{x}^0 + \hat{\mathbf{d}}, \tag{4.10}$$

$$\hat{\mathbf{k}} = \mathbf{C_l}^{-1}\left(\mathbf{A}\hat{\mathbf{d}} + \mathbf{w}^0\right), \tag{4.11}$$

$$\hat{\mathbf{r}} = \mathbf{C_l}\hat{\mathbf{k}}, \tag{4.12}$$

$$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{r}}, \tag{4.13}$$

$$\hat{\sigma}_o{}^2 = \frac{\hat{\mathbf{r}}^T\mathbf{C_l}^{-1}\hat{\mathbf{r}}}{\nu} \tag{4.14}$$

where $\quad \hat{\sigma}_o{}^2 =$ the estimated variance factor and

$\nu =$ the degrees of freedom (number of

observation minus the number of unknowns),

$$\mathbf{C_{\hat{d}}} = \mathbf{C_{\hat{x}}} = \left[\mathbf{A}^T \mathbf{C_l}^{-1} \mathbf{A}\right]^{-1},\tag{4.15}$$

$$\mathbf{C_{\hat{l}}} = \mathbf{A}\left[\mathbf{A}^T \mathbf{C_l}^{-1} \mathbf{A}\right]^{-1} \mathbf{A}^T,\text{ and}\tag{4.16}$$

$$\mathbf{C_{\hat{r}}} = \mathbf{C_l} - \mathbf{C_{\hat{l}}}.\tag{4.17}$$

The matrices $\mathbf{C_{\hat{d}}}$, $\mathbf{C_{\hat{x}}}$, $\mathbf{C_{\hat{l}}}$, and $\mathbf{C_{\hat{r}}}$, are the covariance matrices of the parameter corrections, corrected parameters, adjusted observations, and residuals respectively. The a priori reference factor $\sigma_o^{\,2}$ pertains to the scale of $\mathbf{C_l}^{-1}$. Although the scale of $\mathbf{C_l}^{-1}$ is irrelevant as far as (4.9) and (4.12) are concerned, it is significant in the calculation of the other covariance matrices. If the observation covariance matrix is properly scaled, the estimated variance factor $\hat{\sigma}_o^{\,2}$, will be very close to unity. If $\mathbf{C_{\hat{l}}}$ is unscaled, it may be scaled by dividing by $\sigma_o^{\,2}$.

Each LS iteration begins with the calculation of $\mathbf{w}^0$, the misclosure vector, and $\mathbf{A,}$ the design matrix, for some a priori value of the parameters. This a priori parameter information may be the estimated parameter values $\hat{\mathbf{x}}$ from the previous iteration or some approximation for the first iteration. The remaining unknown quantities are then obtained, in order, from (4.9) through (4.17).

4.2.3 Trilateration Models

4.2.3.1 Circular Trilateration

Consider a cellular telephone with horizontal coordinates (x_c, y_c) and transmitting a message at some unknown time of transmission T_c. The message is received at N cell sites and the TOA at each cell site estimated. Let the coordinates and TOA of the ith cell site be (x_i, y_i) and T_i respectively. The unknown parameters to be estimated are then the telephone coordinates and T_c. The observations are the TOAs at the N cell sites. In

the absence of observational error, the difference between T_c and T_i will equal the geometrical distance between the coordinates (x_c, y_c) and (x_i, y_i) divided by the speed of light, c. When observation errors are present, the preceding is only true with adjusted observations and estimated parameters. In any case, the mathematical model for the ith cell site is

$$f(x,l) = T_i - T_c - \frac{1}{c}\sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} = 0 \text{ .(4.18)}$$

The first design matrix $\mathbf{A}$ is obtained from the partial derivatives of $\mathbf{f(x,l)}$ with respect to the three unknowns, (x_c, y_c) and T_c. For the N cell sites

$$\mathbf{A} = \frac{1}{c}\begin{bmatrix} -\dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_1}{r_1} & -1 \\ -\dfrac{x_c - x_2}{r_2} & -\dfrac{y_c - y_2}{r_2} & -1 \\ \vdots & \vdots & \vdots \\ -\dfrac{x_c - x_N}{r_N} & -\dfrac{y_c - y_N}{r_N} & -1 \end{bmatrix} \qquad (4.19)$$

where $\quad r_i = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2}$.

Because the TOA observations are uncorrelated and unweighted, the unscaled observation covariance matrix $\mathbf{C_l}$ is an identity matrix of dimension N. This form of circular trilateration is often called the pseudoranging mode to distinguish it from the case where the TOT is known and pure ranges, therefore, are measured. Pseudoranges refer to range measurements which include some unknown parameter such as the TOT or a clock offset.

4.2.3.2 Hyperbolic Trilateration with Reference Differencing

As discussed in Chapter 2, the TOT may be eliminated by differencing TOAs between cell sites. Mathematically, this is identical to circular trilateration (pseudoranging mode) as described above. When the correlation in the TOA differences is accounted for, hyperbolic trilateration and circular trilateration (pseudoranging mode) give identical results.

Although the TOT may be solved for in circular trilateration as explained above, it may not be possible to do so in practice. For TOT to be estimated by LS, an approximate TOT must be provided. An approximate TOT of sufficient accuracy may or may not be available depending on the distances involved. For the system presented and tested in this work, a more practical reason for employing TDOA exists. Absolute biases existing at each cell site could not be measured. However, relative biases, or bias differences between cell sites, could be measured. Therefore, only TOA differences could be corrected for the biases and used as observations.

Reference differencing refers to the choice of one TOA observation as a reference and TOA differences formed by differencing that TOA with all others. Let the reference cell site be that corresponding to $i = 1$. The mathematical model for cell sites $i = 2, \ldots, N$ will then be

$$\mathrm{f}(\mathrm{x},1) = \Delta T_{i1} - \frac{1}{c}\sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} + \frac{1}{c}\sqrt{(x_c - x_1)^2 + (y_c - y_1)^2} = 0$$

(4.20)

where $\Delta T_{i1} = T_i - T_1$.

Taking partial derivatives of (4.20) with respect to the unknown parameters (x_c, y_c) results in

$$\mathbf{A} = \frac{1}{c} \begin{bmatrix} -\dfrac{x_c - x_2}{r_2} + \dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_2}{r_2} + \dfrac{y_c - y_1}{r_1} \\ -\dfrac{x_c - x_3}{r_3} + \dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_3}{r_3} + \dfrac{y_c - y_1}{r_1} \\ \vdots & \vdots \\ -\dfrac{x_c - x_N}{r_N} + \dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_N}{r_N} + \dfrac{y_c - y_1}{r_1} \end{bmatrix}. \quad (4.21)$$

Due to the fact that the observations consist of TOA differences, the observations are no longer uncorrelated and the observation covariance matrix is no longer an identity matrix. To account for the correlation between the TOA differences, we write the observation vector as

$$\mathbf{l} = \underbrace{\begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & 0 \\ -1 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ -1 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{H}} \underbrace{\begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_N \end{bmatrix}}_{\mathbf{T}}. \quad (4.22)$$

Note that $\mathbf{l}$ is a column vector of dimension N-1. The unscaled observation covariance matrix is then

$$\mathbf{C_l} = \mathbf{HH}^T = \begin{bmatrix} 2 & 1 & \cdots & 1 \\ 1 & \ddots & 1 & \vdots \\ \vdots & 1 & \ddots & 1 \\ 1 & \cdots & 1 & 2 \end{bmatrix} \quad (4.23)$$

and has dimensions N-1 by N-1, reflecting N-1 independent observations.

4.2.3.3 Hyperbolic Trilateration with Sequential Differencing

In contrast to differencing the TOAs of all cell sites from that of one reference cell site, it is possible to difference the TOAs in a sequential manner. Consider TOA observations for N cell sites. We may form the following N-1 linearly independent TOA differences:

$$
\begin{aligned}
&T_2 - T_1 \\
&T_3 - T_2 \\
&\vdots \\
&T_{N-1} - T_{N-2} \\
&T_N - T_{N-1}.
\end{aligned}
$$

(4.24)

The mathematical model is then

$$
\mathbf{f}(\mathbf{x},\mathbf{l}) = \Delta T_{ii-1} - \frac{1}{c}\sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} +
$$

$$
\frac{1}{c}\sqrt{(x_c - x_{i-1})^2 + (y_c - y_{i-1})^2} = 0
$$

(4.25)

for $i = 2,\ldots,N$ and where $\Delta T_{ii-1} = T_i - T_{i-1}$. The corresponding design matrix is

$$
\mathbf{A} = \frac{1}{c}
\begin{bmatrix}
-\dfrac{x_c - x_2}{r_2} + \dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_2}{r_2} + \dfrac{y_c - y_1}{r_1} \\[2mm]
-\dfrac{x_c - x_3}{r_3} + \dfrac{x_c - x_2}{r_2} & -\dfrac{y_c - y_3}{r_3} + \dfrac{y_c - y_2}{r_2} \\[2mm]
\vdots & \vdots \\[2mm]
-\dfrac{x_c - x_N}{r_N} + \dfrac{x_c - x_{N-1}}{r_{N-1}} & -\dfrac{y_c - y_N}{r_N} + \dfrac{y_c - y_{N-1}}{r_{N-1}}
\end{bmatrix}
$$

.(4.26)

The observation vector is now

$$
\mathbf{l} =
\begin{bmatrix}
-1 & 1 & 0 & 0 & \cdots & 0 \\
0 & -1 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & \cdots & 0 & -1 & 1
\end{bmatrix}
\begin{bmatrix}
T_1 \\
T_2 \\
\vdots \\
T_N
\end{bmatrix}
$$

(4.27)

$$
\mathbf{H} \qquad\qquad \mathbf{T}
$$

but still of dimension N-1. The unscaled N-1 by N-1 observation covariance matrix is

$$\mathbf{C_1} = \mathbf{HH}^T = \begin{bmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \ddots & \vdots \\ 0 & -1 & 2 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{bmatrix}. \qquad (4.28)$$

4.2.3.4 Circular Trilateration with TOA Differences

As discussed in section 4.2.3.2, implementation issues may require the use of TDOA and hence hyperbolic trilateration. One factor which requires the use of TDOA is the effect of SA on GPS time synchronization. In section 2.4.3, the formation of TOA differences was presented as a method of mitigating, if not eliminating, the effect of SA.

However, the use of TOA differences as observations results in observation residuals which also correspond to TOA differences. Residuals are often a useful quantity for testing the quality of the observations. For the application of cellular telephone positioning, residual testing may be used to monitor the integrity of individual cell sites. If the observation residuals correspond to TOA differences, it may not be possible to isolate the cell site causing large residuals.

Therefore, one would ideally wish to use TOA differences as observations in order to derive the implementation benefits described above and at the same time obtain residuals on individual TOAs. This may be accomplished by circular trilateration with TOA differences. Again consider a cellular telephone with horizontal coordinates (x_c, y_c) and transmitting a message at some unknown time of transmission T_c. The message is received at N cell sites, with the coordinates and TOA of the ith cell site being (x_i, y_i) and T_i respectively. We begin by choosing one cell site to be the reference as in the case of hyperbolic trilateration with reference differencing. For the sake of discussion let the

reference site be the first site. The assumption is then made that the TOA at the reference site is equal to the TOT corrected by any bias at the reference cell site, or

$$T_1 = T_c - B_1 \qquad (4.29)$$

where B_1 is the bias of the first cell site. Observations are formed by differencing the TOA at each site with the TOT or, equivalently, the TOA at the reference site. The observation equation for the first cell site will be

$$\Delta T_{11} = T_1 - T_c + B_1 = 0 \qquad (4.30)$$

which is necessarily equal to zero due to the assumption just made. The observation equation for the ith cell site, where $i = 2, \ldots, N$, is

$$\begin{aligned} \Delta T_{i1} &= T_i - T_c + B_i \\ &= T_i - T_1 - B_1 + B_i \\ &= T_i - T_1 + B_{i1}. \end{aligned} \qquad (4.31)$$

Note that in contrast to hyperbolic trilateration in which we obtain N-1 observations from N cell sites, here we have N observations from N cell sites. However, the additional observation is zero and hence there is no new information. What we have essentially are hyperbolic trilateration observations in a circular trilateration type model. That model is

$$f(x,1) = \Delta T_{i1} + \frac{1}{c} R_u - \frac{1}{c} \sqrt{\left(x_c - x_i\right)^2 + \left(y_c - y_i\right)^2} = 0 \qquad (4.32)$$

where $\Delta T_{i1} = T_i - T_1$ (the biases of (4.31) have been disregarded) and R_u is the third unknown, necessary for the circular trilateration mechanism. In (4.18) the three unknowns are (x_c, y_c) and T_c whereas in (4.32) the unknowns are (x_c, y_c) and R_u. From (4.32) it is obvious that R_u is the range from the first (or reference) cell site to the cellular telephone.

Taking partial derivatives of (4.32) with respect to the unknowns the design matrix is found to be

$$\mathbf{A} = \frac{1}{c} \begin{bmatrix} -\dfrac{x_c - x_1}{r_1} & -\dfrac{y_c - y_1}{r_1} & 1 \\ -\dfrac{x_c - x_2}{r_2} & -\dfrac{y_c - y_2}{r_2} & 1 \\ \vdots & \vdots & \vdots \\ -\dfrac{x_c - x_N}{r_N} & -\dfrac{y_c - y_N}{r_N} & 1 \end{bmatrix} \qquad (4.33)$$

which is identical to that of circular trilateration (4.19) with the exception of the sign of the last column. What remains now is to determine the observation covariance matrix. At first glance one might assume that since the observations are TOA differences, the observation covariance matrix must account for this correlation as in (4.23). However, in hyperbolic trilateration any correlation in the measurements will affect the unknowns (x_c, y_c) and must, therefore, be accounted for. In the case under consideration, the correlation in the observations is accounted for by the presence of the third unknown R_u. As a result, the observation covariance matrix is the identity matrix of dimension N.

The advantage of circular trilateration with TOA differences lies in its ability to produce N observation residuals when only N-1 TOA differences are available. Recall that in some circumstances we are forced to use TOA differences because only bias differences are available to correct the observations. In that case, and under hyperbolic trilateration, only residuals on TOA differences are available. Under the formulation just presented,

however, the observation vector is of length *N* since it consists of a zero and *N*-1 TOA differences. The residual vector is, therefore, also of length *N*. That the residual corresponding to the 'zero' observation actually corresponds to the TOA observation of the reference cell site is demonstrated in the results of Chapter 7.

4.2.4 Geometry

A critical factor in the performance of a trilateration positioning system is geometry. The relative positions of the cell sites with respect to one another and the cellular telephone, has a significant impact on positioning accuracy. For methods employing the design matrix **A**, the geometrical strength of a position fix is measured by the Dilution of Precision or DOP (Lachapelle, 1993). The DOP is a unitless number which relates accuracy in the measurement domain to accuracy in the position domain. For the case of position estimated from TOA measurements, this is expressed by

$$\sigma_p = DOP \cdot \sigma_{TOA} \qquad (4.34)$$

where σ_p = the standard deviation of the position estimates, and

σ_{TOA} = the standard deviation of the TOA measurements.

According to (4.34), when the DOP value decreases the amount of error passed from the measurement domain to the position domain, due to geometry, decreases. Consequently, a small DOP value indicates strong geometry and is desirable.

Because DOP is a measure of geometrical strength, it is necessarily a function of the design matrix **A**. Consider the case of circular trilateration for horizontal positioning. From (4.15), the unscaled covariance matrix of the estimated parameters is

$$\mathbf{C}_{\hat{\mathbf{x}}}^{'} = \left[\mathbf{A}^T \mathbf{C_l}^{-1} \mathbf{A} \right]^{-1} = \begin{bmatrix} \sigma_{xx}^{'2} & \sigma_{xy}^{'2} & \sigma_{xt}^{'2} \\ \sigma_{yx}^{'2} & \sigma_{yy}^{'2} & \sigma_{yt}^{'2} \\ \sigma_{tx}^{'2} & \sigma_{ty}^{'2} & \sigma_{tt}^{'2} \end{bmatrix} \tag{4.35}$$

where $'$ denotes an unscaled quantity. For (4.35) to be unscaled requires that $\mathbf{C_l}$ be an identity matrix. Note that $\mathbf{C_l}$ must also be divided by c^2 in order to calculate DOP from (4.35).

The Horizontal Dilution of Precision (HDOP) is a measure of geometrical strength in two dimensions. It is calculated as

$$HDOP = \sqrt{\sigma_{xx}^{'2} + \sigma_{yy}^{'2}} \ . \tag{4.36}$$

Geometrical strength in each of the two individual dimensions may also be quantified. In this work the *x* coordinate refers to the west to east direction (easting) and *y* refers to the south to north direction (northing). Therefore, the DOP for the *x* coordinate is referred to as the East DOP (EDOP) and is given by

$$EDOP = \sqrt{\sigma_{xx}^{'2}} \ . \tag{4.37}$$

In like manner, the North DOP (NDOP) is

$$NDOP = \sqrt{\sigma_{yy}^{'2}} \tag{4.38}$$

and the Time DOP (TDOP), which is the effect of geometry on the TOT estimate, is

$$TDOP = \sqrt{\sigma_{tt}^{'2}} \ . \tag{4.39}$$

In the case of hyperbolic trilateration, equations (4.35) through (4.38) may be used with appropriate $\mathbf{A}$ and $\mathbf{C_l}$ matrices. TDOP is not available since the only two parameters to be estimated are the components of horizontal position. For a given situation, the DOP values will be the same for each of the trilateration types discussed.

4.2.5 Statistical Analysis of Least Squares Results

Following a LS adjustment, it is possible to statistically analyze the results. The LS method does not require the observations to be normally distributed. Statistical testing of results, however, is only valid if the observation errors are normal. In addition, statistical testing also assumes that the LS estimation was performed on a linear model. In the case of a nonlinear model, the last linearized form after iterating must be close to the nonlinear model for the statistical testing to be valid.

The principal purpose of statistical testing, for the present application, is the identification of observations which contain serious errors or blunders. Related to this is the concept of internal reliability. For any statistical testing of the observation residuals, it is imperative that the covariance matrices obtained from LS be correctly scaled. However, prior to discussing these tests, the concept of hypothesis testing is presented.

4.2.5.1 Hypothesis Testing

Integral to statistical testing is the notion of hypothesis testing. A concise overview of this concept is given here. Further details regarding hypothesis testing may be found in Steeves et al. (1987) and Roberts (1993). Consider a random variable some parameter of which LS attempts to estimate. Hypothesis testing has three main components: 1) form a hypothesis regarding the statistical nature of the estimated parameter, 2) choose a test based on a statistic computed from the observations, 3) compare the test statistic with the bounds of a confidence region (Steeves et al., 1987).

To begin then, an initial or null hypothesis is made regarding the parameter of interest. This null hypothesis is denoted by $\mathbf{H_o}$. For instance, we may be interested in estimating the mean of some parameter. The null hypothesis may be that the mean is equal

to a particular value. The validity of this hypothesis is to be tested. There are four possible scenarios. In the event that the hypothesis is correct, the statistical test will lead one to accept or reject the hypothesis. The same is true if the hypothesis is actually false. Table 4.1 summarizes these four outcomes.

Table 4.1   Hypothesis Test Conclusions

	Acceptance of H_o	Rejection of H_o
H_o correct	Correct Conclusion $(1-\alpha)$	Type I Error Committed (α)
H_o incorrect	Type II Error Committed (β)	Correct Conclusion $(1-\beta)$

If H_o is in fact correct, the probability that the hypothesis test will conclude in favor of accepting H_o is equal to $(1-\alpha)$ and is called the confidence level. A type I error is committed when the hypothesis test concludes that H_o be rejected. The probability of this is α, the significance level of the test. The probability of correctly rejecting H_o when it is incorrect, is the power of the test $(1-\beta)$, whereas accepting an incorrect H_o results in a type II error with probability β.

Probabilities α and β are related to one another through the non-centrality parameter, δ_o (Leick, 1995). For every null hypothesis H_o, an infinite number of alternative hypotheses exist. Consider one such alternative hypothesis, H_a. The relationship between H_o, H_a, α, and β is illustrated in Figure 4.1.
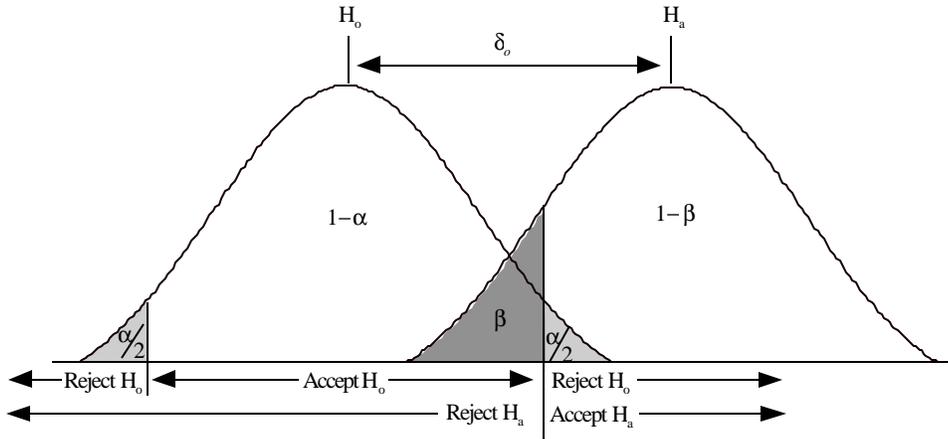
Figure 4.1   Null and Alternative Hypothesis'

In Figure 4.1, a type I error is committed if the statistical test result falls anywhere within the light gray shaded area. A type II error occurs when the test conclusion falls within the dark gray shaded area. Note that for a particular alternative hypothesis and non-centrality factor, it is impossible to reduce both the probability of a type I error and the probability of a type II error to zero. One is minimized at the expense of the other.

4.2.5.2 $\mathbf{c}^2$ Test on the Variance Factor

From equations (4.9), (4.11), and (4.12) we see that the scale of the observation covariance matrix, $\mathbf{C_l}$, does not affect the computation of the estimated parameters or the observation residuals. However, $\mathbf{C_l}$ must be properly scaled to correctly compute the other covariance matrices and the standardized residuals used for statistical testing.

To test whether $\mathbf{C_l}$ is properly scaled, the estimated variance factor, $\hat{\sigma}_o^{\,2}$, is computed from (4.14). If $\mathbf{C_l}$ is properly scaled, both $\sigma_o^{\,2}$ and $\hat{\sigma}_o^{\,2}$ will be unity. The null hypothesis to be tested then is $H_o : \sigma_o^{\,2} = \hat{\sigma}_o^{\,2}$. If the observation residuals $\hat{\mathbf{r}}$ are normally distributed, the test statistic $\nu \hat{\sigma}_o^{\,2} \big/ \sigma_o^{\,2}$ has a χ^2 distribution with ν degrees of

freedom (Steeves et al., 1997). We therefore construct the two-tailed confidence region for σ_o^2 as follows,

$$\frac{\nu \hat{\sigma}_o^2}{\chi^2_{\nu,\alpha/2}} < \sigma_o^2 < \frac{\nu \hat{\sigma}_o^2}{\chi^2_{\nu,1-\alpha/2}}. \qquad (4.40)$$

If σ_o^2 falls within the confidence region of (4.40) one may be (1-α)% certain that it is correct. If it falls outside this interval it is likely that σ_o^2 was incorrectly assumed or that the model used to relate the parameters to the observations is not correct. For example, unmodeled systematic errors in the observations would cause this test to fail. For the application at hand, multipath, which is very difficult to model and account for, is an obvious source of such systematic errors.

4.2.5.3 Statistical Testing of Residuals for Outliers

The detection of any gross errors in the observations is accomplished through an examination of the observation residuals. The null hypothesis to be tested is that the individual estimated residuals, $\hat{r}_i$, are normally distributed with a mean of zero and a standard deviation of $\sigma_{\hat{r}_i}$, where $\sigma_{\hat{r}_i}^2$ is the ith element of the principal diagonal of $\mathbf{C}_{\hat{\mathbf{r}}}$ (4.17). This assumes that σ_o^2 is correctly known. If the observations are normally distributed, the residuals should be as well. The residuals are standardized such that the entire set of residuals is assumed to belong to the normal distribution with zero mean and unit standard deviation, n(0,1). The ith standardized residual is then

$$\hat{r}_i^* = \frac{\hat{r}_i}{\sigma_{\hat{r}_i}}. \qquad (4.41)$$

If the null hypothesis is correct, the standardized residuals should fall within the (1-α)% confidence region or,

$$\hat{r}_i^* = \left| \frac{\hat{r}_i}{\sigma_{\hat{r}_i}} \right| < n(0,1)_{1-\frac{\alpha}{2}}. \qquad (4.42)$$

The statistical testing of (4.42) examines each individual residual in isolation. Of greater interest is the testing of individual residuals "in context". That is we wish to test whether all standardized residuals are simultaneously within a certain confidence interval. To this end, the confidence region of (4.42) is replaced by $1 - \frac{\alpha}{2N}$ where N is the total number of residuals.

If for a particular standardized residual, the test of (4.42) fails, it is likely that the corresponding observation is not normally distributed. The suspected cause of this is that the observation contains a blunder. However, it is possible for the test to fail even in the absence of blunders. Since multipath is not Gaussian, its presence may cause the observations to be non-Gaussian as well. Therefore, the test of (4.42) should be used with caution.

In any case, the results of (4.42) may be used in a number of ways for the cellular positioning application. One option would be to discard that observation and perform another LS adjustment. Observation redundancy must obviously exist in order to do this. Systematic testing of the residuals may also be done to identify a cell site which may be consistently providing poor observations and is, therefore, in need of service.

4.2.5.4 Internal Reliability

Closely associated to the notion of blunder identification is the concept of reliability analysis. Internal reliability is the ability of a system or model to detect a blunder in an observation (Leick, 1995). External reliability refers to an estimate of the effect of an undetected blunder on the estimated parameters.

Reliability may be defined in terms of the controllability of the observations. If the observations are highly controlled, a blunder in one of the observations will be easily detected. A blunder in a poorly controlled observation will not be detected by the statistical testing described in the previous section. How well the observations are controlled is a

function of the redundancy in the observations. Redundancy numbers are defined as the elements of the principal diagonal of the matrix $\left(\mathbf{C}_{\hat{\mathbf{r}}} \mathbf{C}_{\mathbf{l}}^{-1} \right)$. The redundancy number of the ith observation is

$$g_i = \left(\mathbf{C}_{\hat{\mathbf{r}}} \mathbf{C}_{\mathbf{l}}^{-1} \right)_{ii}. \qquad (4.43)$$

Note that due to (4.16) and (4.17), $\mathbf{C}_{\mathbf{l}}$ need not be properly scaled to compute redundancy numbers from (4.43).

An observation is considered to be fully controlled when its redundancy number is unity (Mackenzie, 1985). In that case, all of the observational error appears in the residual for that observation. Should that particular observation contain a blunder, statistical testing of that observation's residuals will detect this. Poorly controlled observations have a redundancy number significantly less than unity. According to (4.43), the residuals of such observations will not reflect any large errors in the observations.

Internal reliability refers to the computation of a marginally detectable error (MDE). The MDE is the smallest error on a particular observation which the model or system will be able to detect. Any observation errors smaller than the MDE will be undetectable. The MDE for the ith observation is

$$\left| \nabla_{oi} \right| = \frac{\delta_o}{\sqrt{g_i}} \sigma_i \qquad (4.44)$$

where σ_i is the standard deviation of the ith observation and $\left| \nabla_{oi} \right|$ refers to the magnitude of the N by 1 vector ∇_{oi} (Roberts, 1993). An important assumption made in the development of (4.44) is the presence of only one blunder at any one time. As a result, there is only one nonzero entry in ∇_{oi} and the magnitude of ∇_{oi} is equal to the MDE for the ith observation.

Due to the presence of the non-centrality factor δ_o in (4.44), there are probabilities attached to the MDE. What in fact internal reliability addresses is the control of type II

error. The null hypothesis here is that there are no blunders in the observations. The alternative hypothesis states that one blunder exists in one of the observations. From (4.44) we may say that there is a $(1-\beta)$ probability that a blunder greater than (4.44) will be detected whereas the probability that a blunder greater than (4.44) remains undetected is β. The probability of committing a type I error, detecting a blunder when in fact none exists, is α.

External reliability estimates the effect of an undetected, and therefore unremoved, blunder on the solution of the LS adjustment. According to the probabilities just discussed, the maximum blunder that one needs to consider is the MDE. The effect on the estimated parameters by a MDE in the ith observation is

$$\nabla\hat{\delta}_{oi} = \left[\mathbf{A}^T\mathbf{C_l}^{-1}\mathbf{A}\right]^{-1}\mathbf{A}^T\mathbf{C_l}^{-1}\nabla_{oi}. \qquad (4.45)$$

Note that the vector of (4.45) shows the effect on the estimated parameters for a MDE on only one observation. Therefore, a total of N vectors like (4.45) may be computed. For an application of reliability to GPS, see Morley (1997).

4.3 Plane Intersection

Plane intersection was introduced by Schmidt (1972) as an alternative to hyperbolic range difference location. Recall that in hyperbolic trilateration, range or TOA differences define hyperbolic lines of position. The TOA difference is a constant for any point on that line of position. Three known stations yield two independent TOA differences and hence two hyperbolas. Where these two hyperbolas intersect is the estimated position fix. Plane intersection uses an approach called LOCA (Location On the Conic Axis). The fundamental theorem of LOCA states that the TOA differences for three stations of known location yield a straight line of position. This straight line is the major axis of a conic. The three stations lie on the conic and the location being estimated lies at one of the foci of the conic. This theorem is illustrated in Figure 4.2 and a proof is given in Schmidt (1972). In hyperbolic trilateration, the stations are located at the foci of the hyperbolas and the location in question

lies on the hyperbola itself. Therefore, LOCA and hyperbolic trilateration may be thought of as duals of each other.
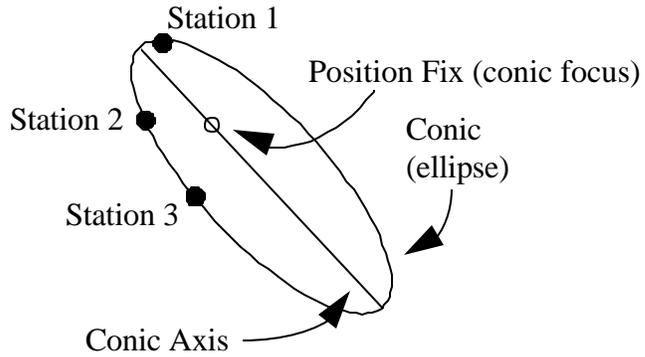
Figure 4.2   Location On the Conic Axis

Once the conic has been defined by the locations of the three stations and their TOA or range differences, it then remains to determine which of the conic's foci corresponds to the correct position fix. If only three stations are available it may be possible to resolve the ambiguity depending on the conic. If the conic is an ellipse, a transmitter at the other focus would generate negative TOA differences compared to a transmitter at the first. In that case the ambiguity may be solved. If the conic is a hyperbola, a transmitter at either of the two foci would generate the same TOA differences and the ambiguity cannot be resolved. In the case of a parabola, one focus has moved off to infinity.

When more then three stations are available, a second conic may be constructed from the triad of stations consisting of two of the original three and the additional fourth. This is illustrated in Figure 4.3. The position fix is clearly the intersection of the two conic's major axes. This method then requires a minimum of four TOAs to estimate an unambiguous position fix.
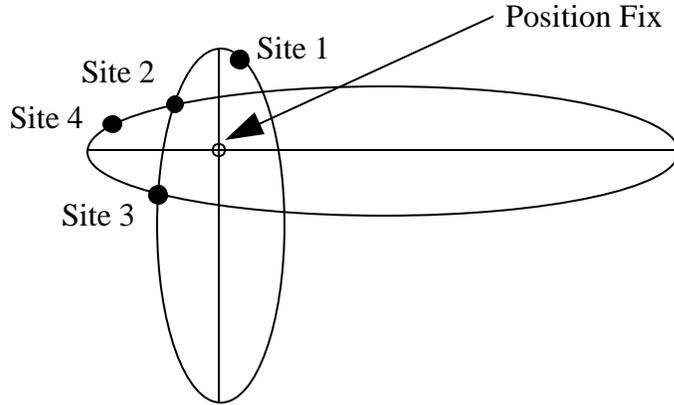
Figure 4.3   Plane Intersection Geometry

To this point the discussion has centered on two dimensional space. The two conics of Figure 4.3 reside in the same plane. In two dimensions at least four stations are required if the position fix is obtained by an intersection of lines. In three dimensions a triad of stations defines a plane instead of a line. Four noncoplanar stations then define two planes which produce a line when they intersect. This line is the axis of a three dimensional conic, an ellipsoid or a circular hyperboloid of two sheets. A second quad of stations will produce a second three dimensional conic axis. The intersection of the two axes occurs where their foci overlap and the result is the position fix. In any case, the intersection of planes provides the position fix and hence the method is called plane intersection.

To formulate the plane intersection algorithm we again consider the case of N cell sites receiving a transmission from a cellular telephone with horizontal coordinates (x_c, y_c). The coordinates and TOA of the ith cell site are (x_i, y_i) and T_i, respectively. Following Schmidt (1996), we form TOA differences for each triad of sites. For example, for the 123rd triad (sites 1, 2, and 3), the following TOA differences are formed:

$$\Delta T_{32} = T_3 - T_2$$
$$\Delta T_{13} = T_1 - T_3 \qquad (4.46)$$
$$\Delta T_{21} = T_2 - T_1.$$

According to the LOCA theorem, there exists a line on which the cell phone resides. This line may be written as

$$Ax + By = D \qquad (4.47)$$

where $A = x_1 \Delta T_{32} + x_2 \Delta T_{13} + x_3 \Delta T_{21}$,

$$B = y_1 \Delta T_{32} + y_2 \Delta T_{13} + y_3 \Delta T_{21},$$

$$D = \frac{\left(\Delta T_{21} \Delta T_{32} \Delta T_{13} + d_1{}^2 \Delta T_{32} + d_2{}^2 \Delta T_{13} + d_3{}^2 \Delta T_{21} \right)}{2} \text{, and}$$

$$d_i{}^2 = x_i{}^2 + y_i{}^2 \quad \text{for } i = 1,\ 2,\ 3.$$

From N cell sites, a total of $\binom{N}{3}$ triads of cell sites are possible where

$$\binom{N}{3} = \frac{N!}{3!(N-3)!}. \qquad (4.48)$$

An equation such as (4.47) is written for each triad and writing these equations in matrix format gives

$$\underbrace{\begin{bmatrix} A_{123} & B_{123} \\ \vdots & \vdots \\ A_{ijk} & B_{ijk} \\ \vdots & \vdots \end{bmatrix}}_{\mathbf{E}} \underbrace{\begin{bmatrix} x_c \\ y_c \end{bmatrix}}_{\mathbf{f}} = \underbrace{\begin{bmatrix} D_{123} \\ \vdots \\ D_{ijk} \\ \vdots \end{bmatrix}}_{\mathbf{g}}. \qquad (4.49)$$

Hence the cellular telephone lies on or is close to each of these lines. The plane intersection solution is then the LS estimate of (x_c, y_c) from (4.49). Note that as (4.49) stands, the LS solution will be weighted where the weights are $\sqrt{A_{ijk}{}^2 + B_{ijk}{}^2}$. To compute the unweighted LS solution to (4.49), each of A_{ijk}, B_{ijk}, and D_{ijk} must first be divided by $\sqrt{A_{ijk}{}^2 + B_{ijk}{}^2}$. When this is done, the residuals being minimized will be the perpendicular error distances to the lines.

It is obvious from (4.49) that there are two unknowns to be solved for, x_c and y_c, and a total of $\binom{N}{3}$ equations. For $N = 3$ there is only one triad and the system of equations (4.49) is under determined. Therefore, at least four cell sites are required. If $N = 4$, there is a total of 4 triads and consequently four equations. However, there exist only two linearly independent triads. This causes the rank of $\begin{bmatrix} \mathbf{E} \vert \mathbf{g} \end{bmatrix}$ to be two, the same as the number of unknowns. As a result, a unique solution to (4.49) is possible.

Equation (4.49) is solved in Matlab$^®$ with the command

$$\mathbf{f} = \mathbf{E} \setminus \mathbf{G}. \tag{4.50}$$

When $\mathbf{E}$ is not square, the Matlab$^®$ operator '\' makes use of Householder reflections to compute the following orthogonal-triangular factorization (Matlab$^®$, 1994b),

$$\mathbf{E} \cdot \mathbf{P} = \mathbf{Q} \cdot \mathbf{R} \tag{4.51}$$

where $\mathbf{P}$ is a permutation matrix, $\mathbf{Q}$ is orthogonal, and $\mathbf{R}$ is upper triangular. The LS solution $\mathbf{f}$ is then

$$\mathbf{f} = \mathbf{P} \cdot \left(\mathbf{R} \setminus \left(\mathbf{Q}^\mathbf{T} \cdot \mathbf{G} \right) \right). \tag{4.52}$$

The error propagation theory of plane intersection is not documented in the literature. In his paper which introduces plane intersection, Schmidt (1972) performs an empirical error analysis. Using simulations, Schmidt determines positional error, due to a particular error in either a TOA difference or a fixed station's coordinates, as a function of the transmitting device location. Unfortunately, the results are of little use for networks of different geometry and distances. A recommended topic for future work is an analytical understanding of how error in the measurement domain is propagated to the position domain.

Note that plane intersection is a closed form algorithm. Approximate values for $\left(x_c, y_c \right)$ are not required. This makes plane intersection a prime candidate for calculating an approximate position which may be used to begin the standardized LS iterative process. For this application, the error propagation theory for plane intersection may be useful for

weighting the approximate parameters in LS. The performance of plane intersection for providing approximate coordinates is investigated in chapters 6, 7, and 8.

CHAPTER 5

TOA ESTIMATION SIMULATIONS

5.1 Introduction

In order to determine the performance of a system such as Cellocate™, simulations were performed in two steps. The first step was to simulate TOA estimation using root MUSIC. This preliminary work was done to verify whether superresolution algorithms such as root MUSIC were capable of sufficient resolution for the application at hand. Once this was verified, positioning simulations were undertaken. The main purpose of the positioning simulations was to determine the positional accuracy of Cellocate™ given the cellular network geometry, multipath, and TOA estimation performance of root MUSIC. The positioning simulations are presented in the next chapter.

This chapter describes the root MUSIC TOA estimation simulations. The simulation of the cellular signal, its modulation, convolution with a channel transfer function, demodulation, and correlation are described in detail. The multipath models used are also presented. The TOA estimation procedure using root MUSIC is then explained. Simulation results are presented in terms of the TOA estimation error mean, RMS, and standard deviation as a function of SNR, multipath delay, phase, and amplitude. The chapter ends with a description of the TOA estimation error models derived from these simulations and used in the positioning simulations.

Software to simulate TOA estimation by root MUSIC was provided by L. Dumont and A. Borsodi. The software simulates the transmission of the 48 symbol preamble of the cellular precursor through a two ray analog cellular telephone channel as well as TOA estimation using root MUSIC.

5.2 Generation and Transmission of the Precursor

The first step in the simulation process was to generate the 48 symbol preamble of the cellular precursor. The 48 symbol preamble consists of a 30 symbol dotting sequence, an 11 symbol Barker sync word and a 7 symbol Digital Colour Code (DCC) in that order. The

dotting sequence is composed of alternating 1's and 0's beginning with a 1. The Barker sync word is 1 1 1 0 0 0 1 0 0 1 0. There are four possible DCCs. Table 5.1 contains all four. For the purposes of these simulations, DCC 2 was used.

Table 5.1   Digital Colour Codes

DCC 0	0	0	0	0	0	0	0
DCC 1	0	0	1	1	1	1	1
DCC 2	1	1	0	0	0	1	1
DCC 3	1	1	1	1	1	0	0

Next, the 48 symbol preamble was mapped to Manchester code. The Manchester code (Haykin, 1989), or split phase code, represents a '0' by a negative pulse followed by a positive pulse. Both pulses are of half-symbol width and have equal amplitudes. A '1' is represented by a positive pulse followed by a negative pulse. The Manchester coded preamble was then upsampled. In AMPS, the data rate is 10 kHz. The Cellocate™ receiver oversamples the signal by a factor of 16. This results in a sampling rate of 160 kHz. When sampled, the 48 symbol preamble becomes 768 samples with amplitude -1, 0, or 1. Samples which occur at a transition are mapped to a 0.

The AMPS standard allows for a data clock offset in the cellular telephone of up to 1 Hz. During the sampling process, distortion in the data clock rate was implemented to determine the effect on TOA estimation. Figure 5.1 shows the TOA estimation error as a function of the data clock frequency. Note that TOA estimation error is expressed in metres. The TOA estimation error in seconds is multiplied by the speed of light to obtain units of metres. This is done since it is often easier to appreciate the significance of a TOA error when it is expressed in units of length. Therefore, in many of the figures to follow, TOA estimation error is expressed in metres.

A high SNR was chosen for Figure 5.1 since the effect of a distorted data clock rate is masked by noise at low SNR. For a data clock offset of 10 Hz, the TOA estimate is in error by 500 m. Fortunately, the data clock offset should not exceed ± 1 Hz, and since it is

common to all TOA estimates, it is canceled when TOA differences are formed for hyperbolic positioning. Therefore, a distorted clock rate was not implemented for the remainder of the simulations.
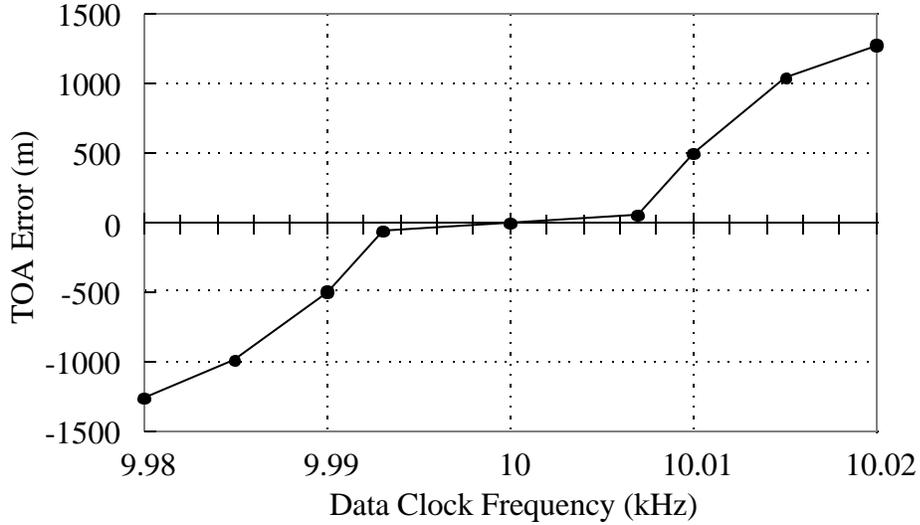
Figure 5.1   TOA Error for Distorted Data Clock Rates (SNR = 30 dB)

After sampling at a rate of 160 kilosamples per second, the preamble signal was integrated in order to ensure a continuous phase in the modulated signal (Proakis, 1989). The signal was then low-pass filtered with a 2nd order Bessel filter. Since the frequency deviation for AMPS is 8 kHz, the cutoff for the transmit Bessel filter was 8 kHz. Following the filter, a complex, equivalent lowpass continuous phase frequency shift keying (CPFSK) signal was generated using the filtered cellular precursor preamble to vary the frequency with a frequency deviation of $\pm$ 8 kHz. The equivalent lowpass signal may therefore be written as

$$v(t) = \exp\left\{ j\left[4\pi T f_d \int_{-\infty}^{t} d(t)dt \right] \right\} \qquad (5.1)$$

where $d(\tau)$ = Manchester encoded, upsampled and lowpass filtered preamble,

T = symbol length in seconds (1/10 kHz), and

f_d = maximum frequency deviation (8 kHz).

5.3 Simulation of the Multipath Channel

The modulated signal was next convolved with the transfer function of the simulated channel. Multipath was simulated using a two ray model. The transfer function of the channel consisted then of the sum of two separate paths. The second path lagged behind the first by an amount corresponding to the type of environment to be simulated. For the purposes of the positioning simulations two propagation environments were assumed - urban and suburban. Turin's empirical data (Turin et al., 1972a) was used to arrive at typical delays for the second path. Turin generated probability densities for the first 10 paths at 1280 MHz for his area B. This area is typical of downtown areas in medium sized cities where, according to Turin's results, LOS is probable. Since Turin does not show probability densities for any of the other areas, this area was chosen to correspond to the general classification of 'suburban'. The excess range of the second path according to Turin's data is approximately 200 feet or 61 m. This corresponds to a time lag of approximately 200 ns which is the value used by the second path in the simulated channel for suburban areas. For 'urban' areas the time lag between the LOS and second path was chosen to be 400 ns.

The relative amplitude of the second arrival with respect to the first was also derived from Turin's data. Morley (1995) traced plots of signal strength as a function of excess range from Turin. From Morley's figures, the second arrival amplitude was found to be 0.4 of that of the LOS path for the 'suburban' area (Turin's area B). For the 'urban' environment (Turin's area A), the ratio was 0.8. The phase of the second arrival with respect to the first was varied from 0° to 315° in steps of 45° to simulate the uniform distribution of received signal phase. The carrier frequency used in the channel transfer response, to account for the phase lag of the second path, was 840 MHz. This corresponds to a wavelength of approximately 36 cm. Table 5.2 summarizes all the parameters pertaining to the urban and suburban channels.

Table 5.2   Simulated 2 Ray Channel Parameters

	Urban Multipath	Suburban Multipath
2nd Arrival Time Lag w.r.t. LOS	400 ns	200 ns
2nd Arrival Amplitude w.r.t LOS	0.8	0.4
2nd Arrival Phase w.r.t. LOS	0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°	
Channel Frequency	840 MHz	

5.4 Demodulation and Correlation

After convolution of the modulated signal with the 2 ray channel transfer function, the FM receiver was simulated. The signal was first filtered with a 2nd order Bessel filter. Since the signal is an equivalent lowpass signal and the bandwidth of the analog cellular channel is 30 kHz, the cutoff frequency of the receive filter was 15 kHz. Average white gaussian noise (AWGN) was then added according to the desired SNR. It will be shown that the performance of MUSIC for TOA estimation is a function of SNR.

Following demodulation and as discussed in Chapter 3, the signal was correlated with a replica of the signal (reference) identically generated except without the channel. Correlation was performed by Fourier transforming both signals to the frequency domain in which one was multiplied by the conjugate of the other and the result inverse Fourier transformed. This resulted in a correlation sequence. For the case of suburban type multipath at 0° phase and an SNR of 15 dB the magnitude of the correlation sequence is shown in Figure 5.2. The number of Fast Fourier Transform (FFT) points used was 1024, the next highest power of 2 over 768 which is the number of samples in the precursor. The peak of the correlation sequence was identified and all data outside of a window centered at the peak was discarded. The limited correlation peak was then Fourier transformed. The size of the window was found to be important for the performance of MUSIC. The number of FFT points used to transform the limited correlation peak to the frequency domain was 64. For the FFT result to be accurate, the windowed data should be periodic and the number of

FFT points should be equal to the data window or some multiple of it. Since there are 16 samples per symbol, the correlation peak should be 32 samples wide. This suggests a 32 sample data window. However, by trial and error it was found that a 15 sample data window gave much better performance at lower SNRs. Sixty-four FFT points is still an approximate multiple of the 15 point data window.
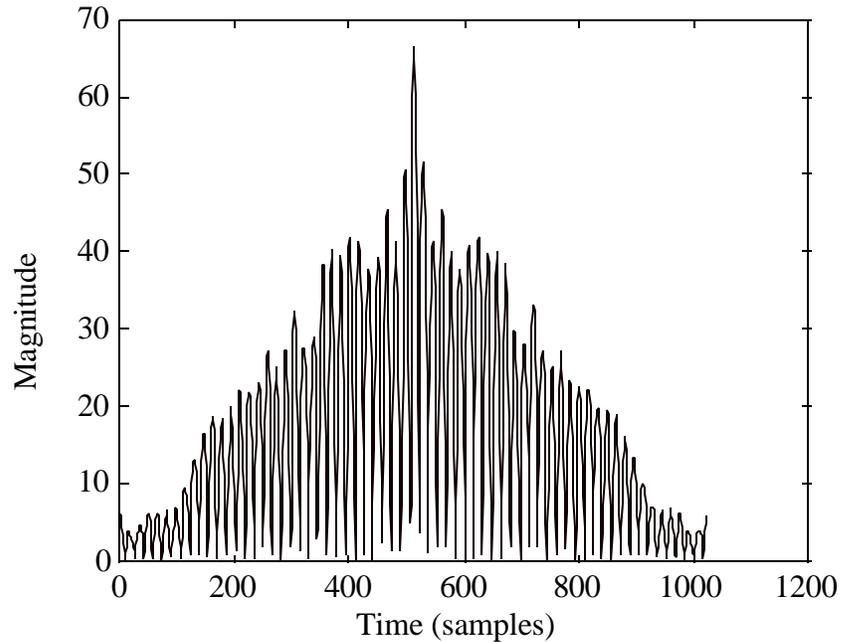
Figure 5.2   Correlation Sequence for Suburban Channel with 0° Multipath Phase

Next, the DC level in the finite length correlation sequence was matched to that of the reference correlation sequence. As discussed earlier, the frequency representation of the original correlation function (reference) must be divided from that of the received signal correlation function in order to yield point sources in noise. Different DC levels in the two correlation functions was found to impact the performance of MUSIC. Therefore, an

iterative process was used to shift the DC level of the received correlation sequence to match the level of the reference correlation sequence.

A problem arises, however, due to the time lag assigned to the LOS path in the channel model. The 2-ray channel model may be written as

$$h(t) = a_1 e^{-j2\pi f_c t_1} + a_2 e^{-j\left(2\pi f_c t_2 + f_2\right)} \qquad (5.2)$$

where a_1, a_2 = amplitudes of the LOS and 2nd ray respectively,

t_1, t_2 = time lags of the LOS and 2nd ray respectively,

f_c = carrier frequency, and

f_2 = phase lag of 2nd path with respect to LOS.

As previously mentioned, the reference correlation pulse does not include the effect of the channel model. Any nonzero value of t_1 in equation (5.2) introduces a phase shift in the received signal. As a result, the phase of the received correlation sequence will not match that of the reference correlation sequence which assumes t_1 to be zero. When the DC level of the received correlation sequence is adjusted to match that of the reference correlation, the phase is affected. Simulations were conducted at each particular SNR by varying the value of t_1 from -0.5 of a sample to 0.5 of a sample since a signal may arrive anywhere within one sample. When $t_1 = 0$, the assumption holds and adjusting the DC level actually improves accuracy. When t_1 is nonzero, adjusting the DC level of the received correlation to match that of the reference affects the phase such that accuracy suffers. Therefore, DC level matching was not employed in the simulations.

Figure 5.3 illustrates the degree to which DC level adjustment does improve accuracy when $t_1 = 0$. Plotted in Figure 5.3 is the difference, in RMS TOA error, between the case when the correlation DC level is adjusted and the case when it is not adjusted. The positive differences correspond to a reduction in the RMS error when the DC level is adjusted. The three curves illustrate the differences for urban and suburban levels of multipath (parameters of Table 5.2 and phase of 0°) as well as the case of no multipath. At higher SNRs, a multipath floor exists. As noise no longer becomes significant, multipath causes a fixed TOA estimation error. The result is a fixed RMS error difference as seen in

73

the urban and suburban curves of Figure 5.3. However, in the absence of multipath, the improvement due to DC level adjustment continues to be a function of SNR. With no noise present, DC level adjustment would be meaningless since the received correlation sequence and reference correlation sequence would be equivalent.

At an SNR of 10 dB in Figure 5.3, the improvement is about 5 m or less. As will be seen later, the RMS TOA estimation error at this SNR is approximately 100 m.   The improvement is therefore 5% or less. At higher SNRs the improvement is also on the order of 5%.
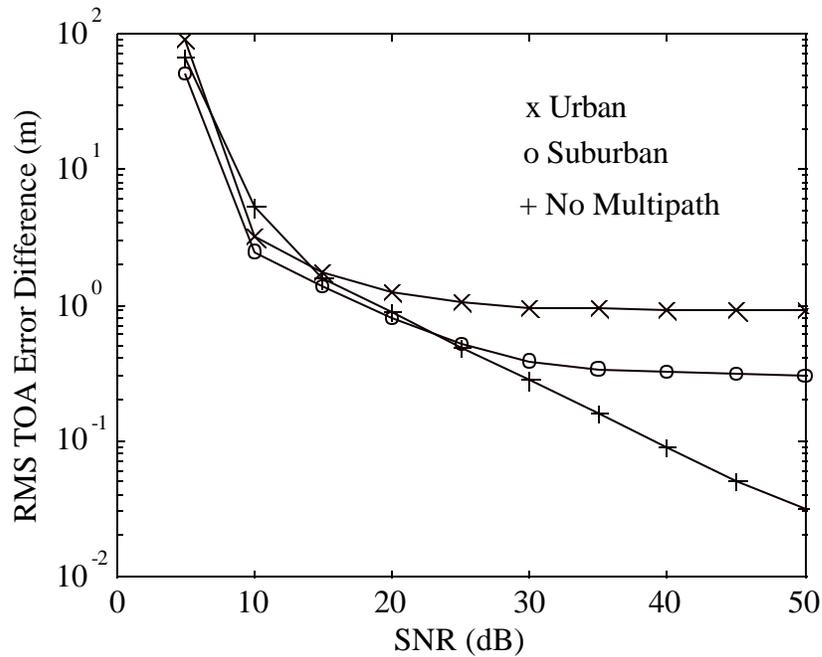
Figure 5.3   Difference in RMS TOA Error between Adjustment of Correlation   DC Level and No Adjustment of Correlation DC Level

Note the linear behavior of the curve corresponding to no multipath, for SNRs greater than 10 dB. At 10 dB, the curve 'breaks' due to the FM threshold effect (Haykin, 1989). Below an SNR of 10 dB, the FM receiver output SNR is no longer a linear function

of the input SNR. The result is impulse-like components in the receiver output and an increase in the TOA estimation error. The FM threshold effect is visible in many of the following figures.

An underlying assumption of MUSIC in its usual application of frequency domain estimation is that the signals to be estimated are uncorrelated complex sinusoids in white noise (Haykin, 1991). When MUSIC is used in the opposite direction of estimating time domain signals from frequency domain data, the same assumption must hold. Therefore, it is necessary that the time domain data be modeled as point sources in white noise. To achieve this, the effects of the lowpass filters and the correlation process must be removed from the frequency domain data before the application of MUSIC. This deconvolution is performed by dividing from the frequency domain data the frequency representation of an identical signal (reference signal) to that received with the exception that it has not passed through the channel. Since it is the impulse response of the channel which is being estimated, the effect of the channel must necessarily remain in the signal.

Both the received frequency domain data and the reference signal were bandlimited before deconvolution. This is necessary due to the nulls in the sinc^2 function at $\pm$ 10 kHz, the data rate. Dividing by the sinc^2 function causes noise enhancement at the nulls. Both signals were, therefore, bandlimited to $\pm$ 10 kHz. For 64 FFT points and a sampling frequency of 160 kilosamples per second, this corresponds to discarding all frequency domain data outside of $\pm$ 4 FFT points of 0 Hz. This parameter was varied and a value of 4 was found by trial and error to give the best performance.

At this point the frequency domain data was ready for processing by root MUSIC. The frequency domain data was first organized into a data matrix, the form of which implements forward and backward smoothing (Haykin, 1991). As discussed in Chapter 3, forward and backward smoothing are effective methods of decorrelating correlated signals. The bandlimited frequency domain signal consists of 9 points (0 Hz point $\pm$ 4 points) and may be written as

$$\mathbf{x} = \begin{bmatrix} x_1 x_2 \cdots x_9 \end{bmatrix}^T \tag{5.3}$$

where T indicates transpose. To implement smoothing, the data vector is partitioned into subarrays. The optimal subarray length is approximately 0.75 times the length of the data vector **x** (Tufts and Kumaresan, 1982). In this case the subarray length was set to 7. The data matrix then consists of two halves, those subarrays arranged in a forward direction and the same subarrays conjugated and ordered in the reverse direction. The data matrix **A** is

$$\mathbf{A}^H = \begin{bmatrix} x_7 & x_8 & x_9 & x^*_1 & x^*_2 & x^*_3 \\ x_6 & x_7 & x_8 & x^*_2 & x^*_3 & x^*_4 \\ x_5 & x_6 & x_7 & x^*_3 & x^*_4 & x^*_5 \\ x_4 & x_5 & x_6 & x^*_4 & x^*_5 & x^*_6 \\ x_3 & x_4 & x_5 & x^*_5 & x^*_6 & x^*_7 \\ x_2 & x_3 & x_4 & x^*_6 & x^*_7 & x^*_8 \\ x_1 & x_2 & x_3 & x^*_7 & x^*_8 & x^*_9 \end{bmatrix}. \qquad (5.4)$$

Once the data matrix has been constructed, a singular value decomposition (SVD) is performed on **A**. The resulting squares of the singular values and the right singular vectors are the same as the eigenvalues and eigenvectors of the correlation matrix

$$\Phi = \mathbf{A}^H \mathbf{A}. \qquad (5.5)$$

The SVD of **A** produces $M+1=7$ singular values where $M+1$ is the subarray length. These singular values are divided into two groups according to a threshold. The threshold used was the largest singular value divided by $2*(N-M+1)$ where N is the total number of points in **x**. All singular values, and their associated right singular vectors, which exceed the threshold correspond to the signal subspace. Those singular values, and their corresponding right singular vectors, which do not exceed the threshold, correspond to the noise subspace. It is these singular vectors which span the noise subspace. Any steering vector which corresponds to a signal will be orthogonal to that space. In the event that no singular values exceed the threshold, a default number of 'signal' singular values was used. That default number is $2*(N-M)$ which yields 6 in this case. Yamada et al. (1991) report that in terms of signal estimation accuracy, it is better to overestimate rather than underestimate the

76

number of signals L. Since the subarray size is 7, the total number of singular values is also 7. Therefore, in order to have at least one singular right vector to span the noise subspace, L may be overestimated to a maximum of 6.

A matrix is formed from the right singular vectors corresponding to the $M+1-L$ smallest singular values. Let the right singular vectors be $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_{M+1}$. The order of these singular vectors is such that the corresponding singular values $\sigma_1, \sigma_2, \cdots \sigma_{M+1}$ are arranged in descending order. The matrix spanning the noise subspace is then

$$\mathbf{V}_{noise} = \begin{bmatrix} \mathbf{v}_{L+1} & \mathbf{v}_{L+2} & \cdots & \mathbf{v}_{M+1} \end{bmatrix} \tag{5.6}$$

where each singular vector is of dimension $(M+1) \times 1$. It follows that the dimensions of $\mathbf{V}_{noise}$ are $(M+1) \times (M+1-L)$.

In root MUSIC, the MUSIC spectrum is represented by

$$\hat{S}_{\text{MUSIC}}(z) = \frac{1}{D(z)} . \tag{5.7}$$

The polynomial $D(z)$ is created from $\mathbf{s}^H(t)\mathbf{V}_{noise}\mathbf{V}_{noise}^H \mathbf{s}(t)$, the MUSIC spectrum denominator, by substituting the variable e^{jt} with z. The coefficients of the polynomial are equal to the sums of the diagonals of the matrix $\mathbf{V}_{noise}\mathbf{V}_{noise}^H$. Those roots of the polynomial $D(z)$ which are closest to the unit circle should correspond to signals and the phase of those roots to the normalized time of those signals. It is not necessary, however, to consider all roots as possible signal candidates. When the received correlation sequence is divided by the reference correlation sequence, the phases are differenced. Ideally the phases are identical and the resulting phase is zero. Consequently, only those roots with phase within some sector centered at zero need be considered. In this case the sector width was -0.5 symbols (-45° or -8 samples) to 0.0625 symbols (5.625° or 1 sample). Since multipath signals always arrive after the LOS signal, the negative sector bound was made larger than the positive. Should the most dominant arrival be multipath and not LOS, it is desirable to consider roots prior to zero phase and not after. Even when the dominant arrival is LOS, it is expected to be close to zero and not significantly after.

That root within the sector and closest to the unit circle was chosen. Its phase in radians was then converted to samples by

$$delay = \frac{phase}{2\pi / NFFT}$$
(5.8)

where *NFFT* is the number of FFT points (in this case 64). This delay is root MUSIC's estimate of the LOS arrival with respect to the peak identified after the correlation process. The correlation peak itself must be referenced to the true TOA of the LOS path in order to calculate the error in TOA. When the time lag for the LOS signal is zero and the entire precursor is used for correlation, the correlation peak will occur at *NFFT* / 2 + 1 samples with respect to zero time. Note that the number of FFT points used in the correlation process (and therefore of significance here) is 1024 whereas the number of FFT points used in the superresolution process is 64. The TOA estimate of root MUSIC is then

$$T\hat{OA}_{MUSIC} = \left(delay + peak - \left(\frac{NFFT}{2} + 1 \right) \right)$$
(5.9)

where $T\hat{OA}_{MUSIC}$ = root MUSIC TOA estimate in samples,

$delay$ = root MUSIC estimate with respect to correlation peak in samples,

$peak$ = number of samples from start of correlation sequence to the peak,

$NFFT$ = number of FFT points used for correlation (1024).

The error in the root MUSIC TOA estimate, expressed in metres, is calculated by

$$TOA_{error} = \left(T\hat{OA}_{MUSIC} - LOS\ lag \right) * c / \left(R * f_{data} \right)$$
(5.10)

where $LOS\ lag$ = time lag of the LOS path in samples,

c = propagation speed (speed of light = 299792.5 m/s),

R = oversampling factor (16), and

f_{data} = data rate (10 kHz).

.5.5 Cramer-Rao Lower Bound

It is possible to compute the Cramer-Rao lower bound (CRLB) for TOA estimation. The CRLB is a lower bound on the variance of any unbiased estimator. It states that the variance of any unbiased estimator $\hat{\theta}$ must satisfy (Kay, 1993)

$$\text{var}\left(\hat{\theta}\right) \geq \frac{1}{-E\left[\dfrac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial \theta^2}\right]} \tag{5.11}$$

where E is the expectation operator, and

$p(\mathbf{x};\theta)$ is the probability density function of the observed

sample

set $\mathbf{x}$.

When estimating a parameter of a signal in white Gaussian noise, equation (5.11) becomes

$$\text{var}\left(\hat{\theta}\right) \geq \frac{1}{\displaystyle\sum_n \left(\dfrac{\partial s[n;\theta]}{\partial \theta}\right)^2} \tag{5.12}$$

where $s[n;\theta]$ is the nth sample of the signal s with parameter θ.

For the case of TOA estimation in the absence of multipath, the received signal can be written as

$$x(t) = s(t - \tau_o) + w(t) \tag{5.13}$$

where $s(t)$ is the transmitted signal,

τ_o is the propagation time from transmitter to receiver, and

$w(t)$ is white Gaussian noise.

Using equation (5.12) it can be shown (see (Kay, 1993) for details) that for the signal model (5.13), the CRLB for range estimation ($\hat{R}$), is

$$\operatorname{var}\left(\hat{R}\right) \geq \frac{c^2}{2 \cdot SNR \cdot \overline{F^2}} \qquad (5.14)$$

where c = the speed of light,

$$\overline{F^2} = \frac{\int_{-\infty}^{\infty} (2\pi F)^2 |S(F)|^2 dF}{\int_{-\infty}^{\infty} |S(F)|^2 dF} \quad \text{and,} (5.15)$$

$S(F)$ is the Fourier transform of $s(t)$.

This is equivalent to the CRLB for TOA estimation expressed in metres. Since the expression $\overline{F^2}$ is a measure of the bandwidth of the signal, it is clear that the CRLB is inversely proportional to the SNR and the signal bandwidth.

5.6 Simulation Results

TOA estimation simulations were carried out for various levels of multipath delay (including the urban and suburban levels given in Table 5.1). Other parameters varied were the amplitude of the reflected signal, its phase with respect to the LOS signal and SNR. To obtain TOA error statistics for any combination of errors, the time lag of the LOS path was varied from -0.5 of a sample to +0.5 of a sample in increments of 0.002. The mean, standard deviation and RMS of the TOA estimation errors are, therefore, based on 500 trials.

The first set of results, shown in Figures 5.4 and 5.5, illustrate the effect of multipath delay. Figure 5.4 plots the mean TOA error, expressed in metres, as a function of SNR. Figure 5.5 contains the RMS of the TOA estimation error, also expressed in metres. In all cases, the phase of the multipath ray is $0°$ with respect to the LOS ray and the amplitude of the multipath ray is 0.5 of that of the LOS ray. Beside each curve appears the time lag of the multipath ray with respect to the LOS ray. Figure 5.5 also shows the CRLB on the standard deviation of the TOA estimate.
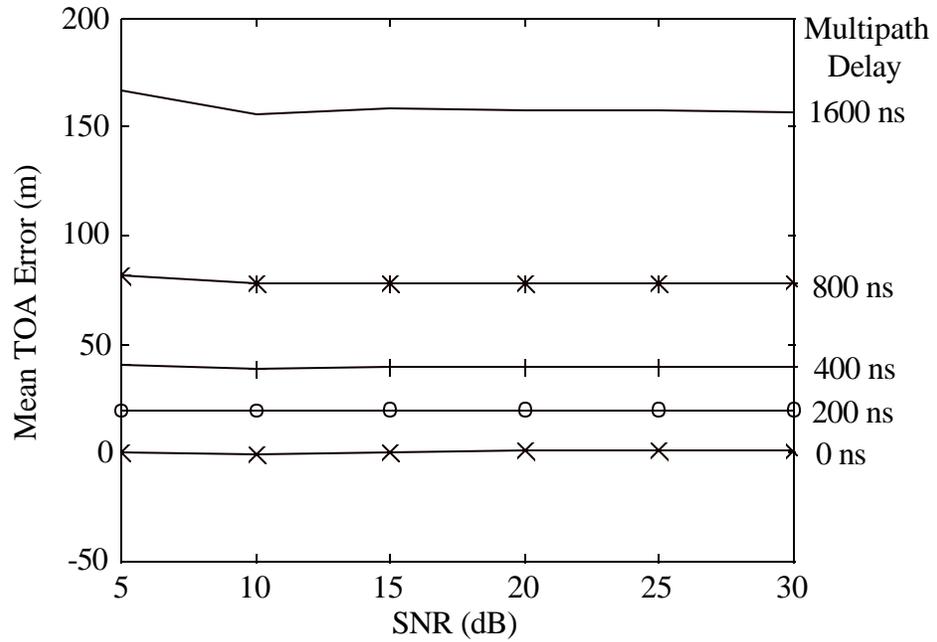
Figure 5.4   Mean TOA Estimation Error for Various Levels of

Multipath Delay (Multipath Phase = 0°, Multipath Ampl. = 0.5)

The standard deviation of the TOA estimation error is approximately equal for all five levels of multipath delay in Figures 5.4 and 5.5 and was not plotted. The standard deviation curve for each of the five multipath levels is virtually identical to the RMS curve for 0 ns multipath delay, since for that level of multipath the mean error is zero.
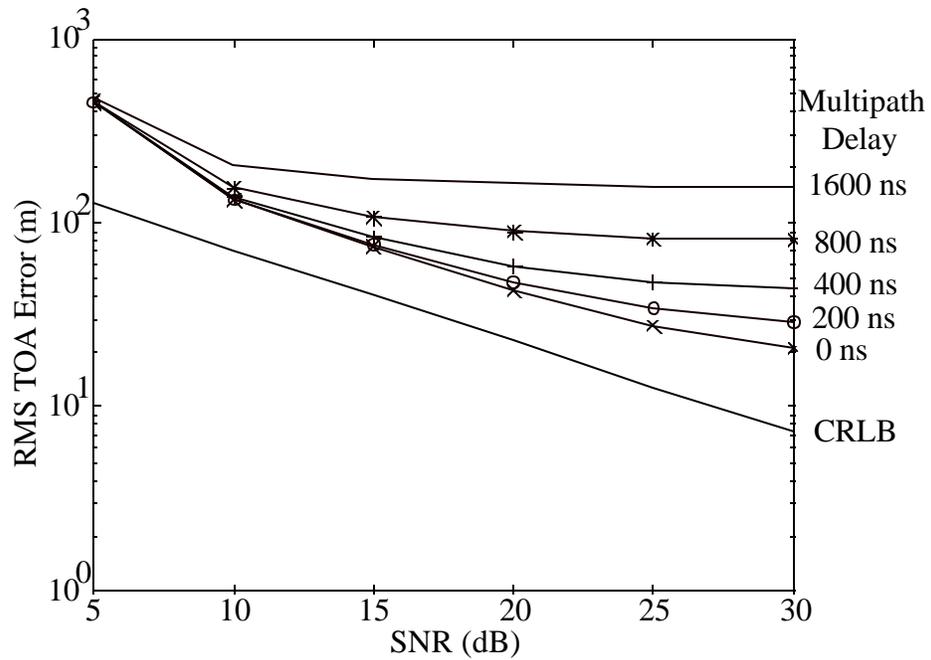
Figure 5.5   RMS TOA Estimation Error for Various Levels of
                 Multipath Delay (Multipath Phase = 0°, Multipath Ampl. = 0.5)

It is clear from these results that the bias in root MUSIC TOA estimation is a function of the multipath delay spread whereas the standard deviation is a function of SNR. Due to the FM threshold effect below 10 dB, the differences in bias are insignificant compared to the standard deviation and therefore the RMS curves converge. At higher SNR, where the standard deviation is on the order of the biases, the RMS curves are distinct. The TOA estimation errors were calculated by subtracting the true TOA from the root MUSIC TOA estimate. As expected, the bias due to multipath is positive. It is interesting to note that the bias in seconds due to multipath is approximately one third of the multipath delay.

The phase of the multipath ray with respect to the LOS ray also has a significant effect on the TOA estimation performance of MUSIC. This is especially true when the two arrivals are out of phase. Then, fading due to multipath will exacerbate an already low SNR.

To demonstrate this, in Figures 5.6 and 5.7 respectively, the mean and standard deviation of TOA estimation error are plotted versus SNR for two rays in phase as well as 90° and 180° out of phase. The delay of the multipath ray with respect to the LOS is 400 ns and the amplitude is 0.8 of that of the LOS ray. These parameters correspond to the urban environment parameters of Table 5.2.

Note that the absolute value of the mean TOA error is plotted. This was necessary only because of the 180° data. Both the 0° and 90° mean data are positive as expected. The 180° data, however, are negative. Since negative data cannot be plotted on a log axis the absolute value was taken.
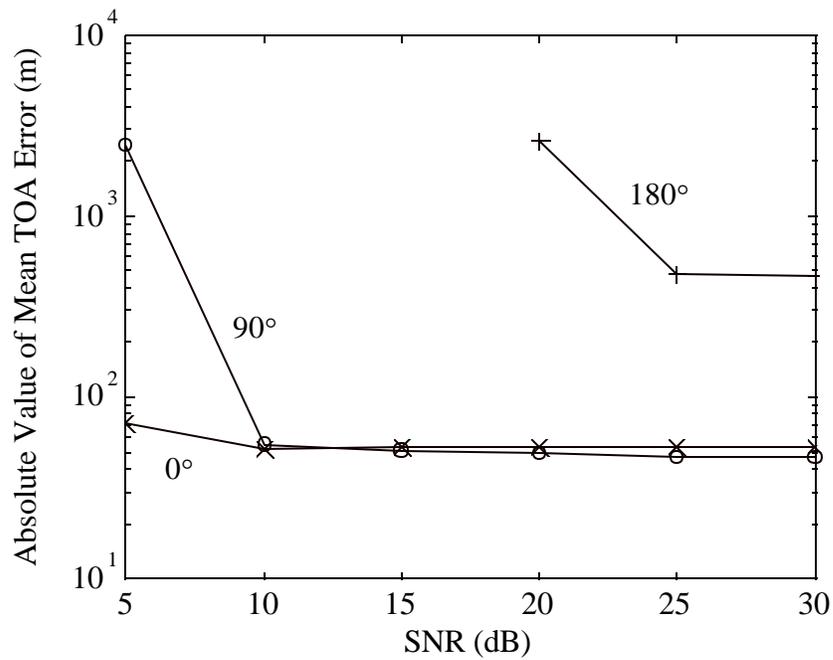
Figure 5.6   Mean TOA Estimation Error as a Function of Multipath Phase
(Multipath Delay Spread = 400 ns, Multipath Ampl. = 0.8)

When the two rays are in phase, the mean error is approximately 50 metres which generally agrees with the data of Figure 5.4. For a 90° phase difference at SNRs above 10 dB the mean error is also in the order of 40 to 50 metres. Below 10 dB, however, the effect

of fading begins to reveal itself. This is most evident in the 180° curve. Fading effectively lowers the SNR. The SNR was calculated as the ratio of the power of the transmitted signal to the power of the noise. The 180° out of phase signal destructively interferes with the LOS signal - the result being a combined signal of reduced power. In this case the amplitude of the multipath ray is only 20% less than that of the LOS ray causing a large amount of fading. Below 20 dB and for the case of 180° phase, the combined effect of fading and an increased noise level is a meaningless correlation sequence. The peak was so far in error that part of the required data window lay outside the correlation sequence and no solution could be generated. Above 20 dB a solution could be generated and at 30 dB the standard deviation is just over 100 metres. However, confidence in the solution is low due to the magnitude of the mean error and the fact that it is negative.
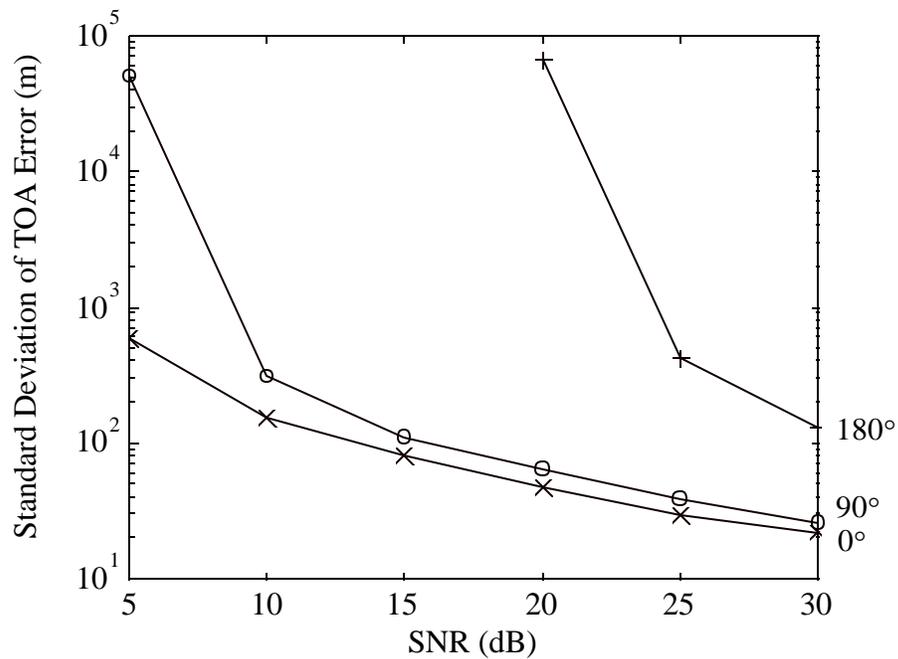
Figure 5.7   Standard Deviation of TOA Estimation Error as a Function of Multipath Phase (Multipath Delay Spread = 400 ns, Multipath Ampl. = 0.8)

The standard deviation curves of Figure 5.7 also demonstrate the effect of fading. Again, there is little difference between 0° and 90° at higher SNRs whereas below 15 dB fading becomes significant. For 180°, the effect of destructive interference is obvious.

The amplitude of the multipath arrival with respect to the LOS arrival is also important. The mean and standard deviation plots of Figures 5.8 and 5.9 illustrate this. The mean errors in the case of a multipath phase of 180° are again negative and the absolute values are plotted. In each of the figures six plots are shown. Three curves correspond to the case of inphase multipath with amplitudes of 0.2, 0.5, and 0.8 of that of the LOS arrival. The other three correspond to 180° multipath with the same amplitudes. In terms of the mean TOA error, the effect of increased multipath amplitude is an increase in the TOA estimation bias. For the 0° phase, the increase in bias, as amplitude increases from 0.2 to 0.8, is approximately 35 m. The effect of amplitude is far more dramatic in the case of 180° multipath. When destructive interference is small (0.2 amplitude) there is little difference between the 0° and 180° multipath TOA estimation means apart from the sign. However, there is a difference of approximately 400 m between the 0.8 amplitude curves.
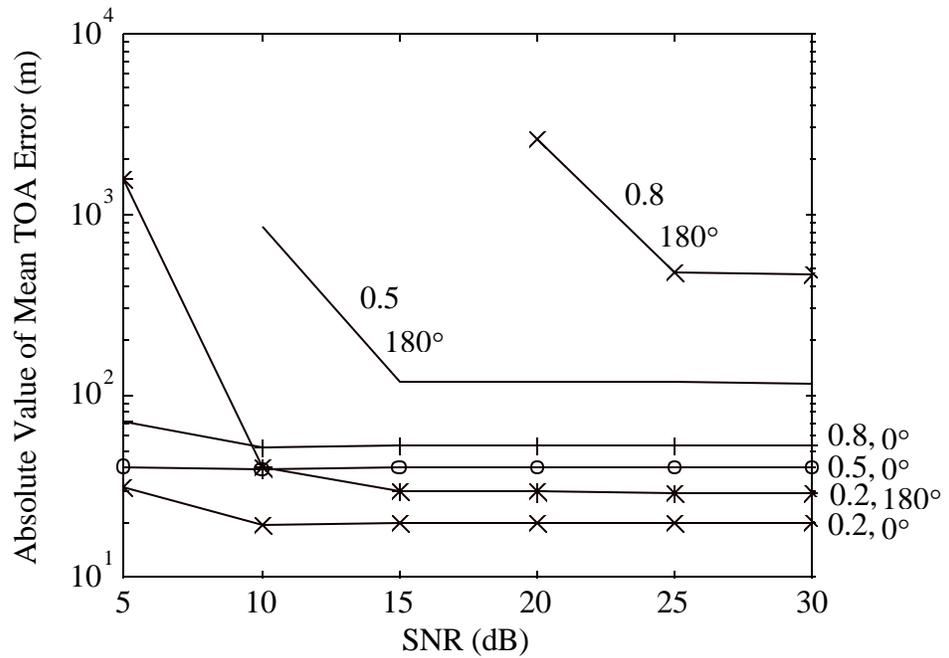
Figure 5.8   Mean TOA Estimation Error as a Function of Multipath Amplitude   and Phase (Multipath Delay Spread = 400 ns)

The fact that amplitude is more significant in the case of out of phase multipath can also be concluded from the standard deviation curves of Figure 5.9. For inphase multipath there is little difference in standard deviation as the multipath amplitude changes. The opposite is true in the case of 180° multipath.
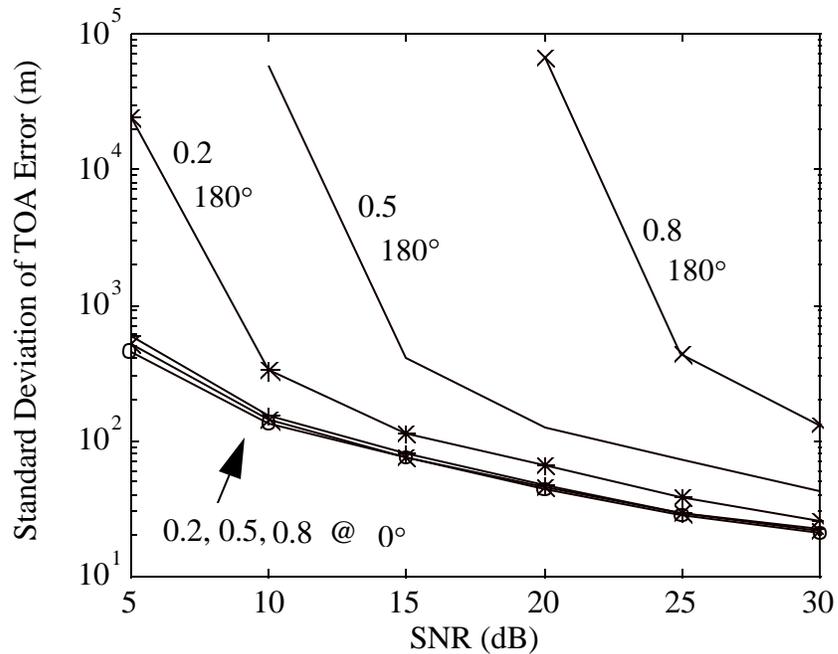
Figure 5.9   Standard Deviation of TOA Estimation Error as a Function of
              Multipath Amplitude and Phase (Multipath Delay Spread = 400

ns)

To conclude, these simulations verify that TOA estimation by root MUSIC is subject to multipath as expected. Out of phase multipath causes destructive interference and increases TOA estimation error. Low SNR and high multipath amplitudes worsen the situation. TOA estimation error also increases as the time lag between the LOS and multipath arrivals increases. One would expect, however, that the error would begin to decrease as the time lag continues to increase. For example, if root MUSIC could only meet the Rayleigh resolution criteria, the multipath arrival should have no effect on TOA estimation of the LOS arrival when the lag between them is the period of one symbol (100 000 ns). However, a multipath delay of this magnitude is unrealistic for the application at hand. Realistic delay spreads for cellular signals as determined by Turin (1972a) are approximately 1/200 of a symbol period and the largest investigated here was 1/62. Dumont

(1994) investigated delay spreads of 1/4 of a chip (or symbol) to 1/32 of a chip and found that the mean TOA error increases as delay spread decreases, the opposite of the findings here. It may be inferred then, that the delay spread threshold which corresponds to a maximum TOA estimation error lies between 1/32 of a symbol and 1/62 of a symbol. In any case, this delay spread threshold is much higher than what could normally be anticipated in the cellular propagation environment. Therefore, for this application, an increase in delay spread will cause an increase in TOA estimation error.

5.7 Models of MUSIC TOA Estimation Error

For the purposes of including MUSIC TOA estimation error in the positioning simulations to follow, models were created. The purpose of these models was to add MUSIC TOA estimation error to the range between a simulated cellular telephone position and a cell site. As discussed in section 5.3, two propagation environments are assumed - urban and suburban. The values of the multipath parameters for the two environments are given in Table 5.2.

Simulations were performed for each of the two environments at all the phases listed in Table 5.2. As before, the time lag of the LOS path was varied from -0.5 samples to 0.5 samples but in increments of 0.005. The statistics at each SNR are therefore based on 200 samples. The SNR was varied from 12 dB to 52 dB in steps of 4 dB. As will be discussed later, the minimum allowed SNR in the positioning simulations is 13 dB and as the following results show, both the mean and standard deviation are essentially constant above 50 dB.

The TOA estimation error mean and standard deviation curves were modeled using polynomials. Polynomials were chosen for two main reasons. The first is the convenience of the polynomial functions contained in Matlab[®], the language in which all the simulation software was written. Matlab[®] includes functions to fit a polynomial to some data and to evaluate the polynomial at any parameter value. Secondly, polynomials were found to fit the curves with an accuracy sufficient for the application. In almost all cases a polynomial order

between five and eight was sufficient to ensure the model matched the simulation values to within 1 m.

Figures 5.10a and 5.10b contain the mean and standard deviation of MUSIC TOA estimation errors for the urban environment. Each figure contains eight plots - the TOA error mean and standard deviation for four different multipath phases. Each plot is labeled with both the multipath phase and the order of the polynomial used to fit the simulation data. Within each plot appear eleven small circles. These correspond to the simulation data. The continuous line is the polynomial model evaluated from 13 dB to 52 dB in steps of 0.01 dB.

All urban simulation points were used for polynomial fitting except in the cases of 180° and 225° multipath phase. In the case of 180° the first three simulation points of the mean (16 dB, 20 dB, and 24 dB) were not used for modeling (as discussed earlier, no TOA solution could be generated at 12 dB and a phase of 180° due to low SNR and fading). Similarly, the first two points of the standard deviation were not used. At low SNR and a high degree of fading, the TOA estimation process is extremely difficult. As a result, it was very difficult to obtain a statistically valid sample set. It was found that even with 200 samples, the mean would significantly change at low SNR when the seed of the random generator, used to produce the noise, was changed. At high SNR and for less destructive multipath, the mean TOA estimation error was the same for a number of random number generator seeds. The conclusion drawn from this is that the mean TOA estimation error shown in the 180° plot of Figure 5.10b is only one realization - particularly at SNRs lower than 24 dB. The first three points were, therefore, not used to model the mean. For the same reason the first two points (16 dB and 20 dB) were not used to model the standard deviation. The same can be said in the case of urban multipath of phase 225°. In this case only the simulation point at 12 dB was not used in modeling the mean. All points were used to model the standard deviation since there are no anomalies in the standard deviation data.

The simulation data and models for the suburban environment are given in Figures 5.11a and 5.11b. All points were used for modeling except for the first (12 dB) in the case of the mean TOA error for 180° multipath. Again this is due to a lack of confidence in the simulation results at an SNR of 12 dB.
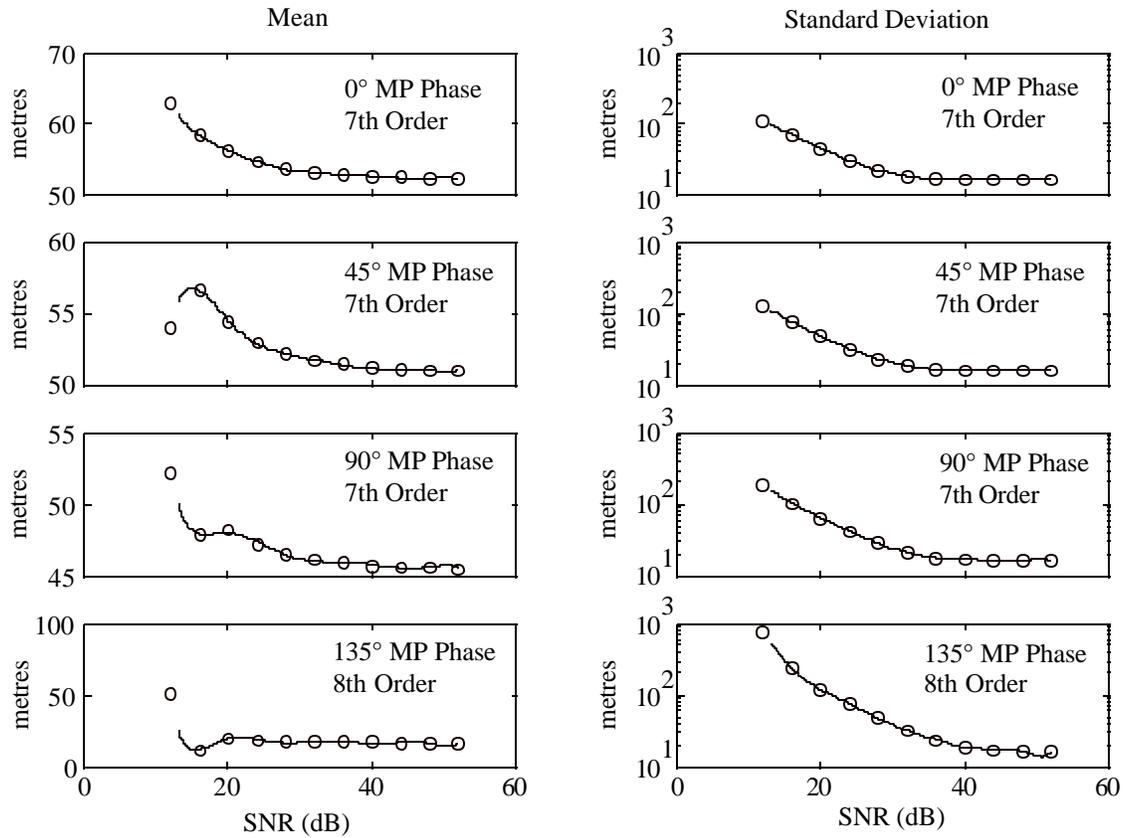
Figure 5.10a   MUSIC TOA Estimation Error Models (Urban)

('o' simulation points, '-' polynomial model)

In the positioning simulations to follow, the models of Figure 5.10 and 5.11 were used to add TOA estimation error to simulated multipath corrupted ranges. The appropriate model - urban or suburban - was chosen based on the location of the simulated cellular telephone position.   The multipath phase, with respect to that of LOS, was randomly generated from a uniform distribution with outcomes {0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°}. The polynomial models for the particular environment and phase were then used to

generate the mean and standard deviation of the TOA estimation error. The calculated SNR
of the simulated received signal served as the parameter for the models.
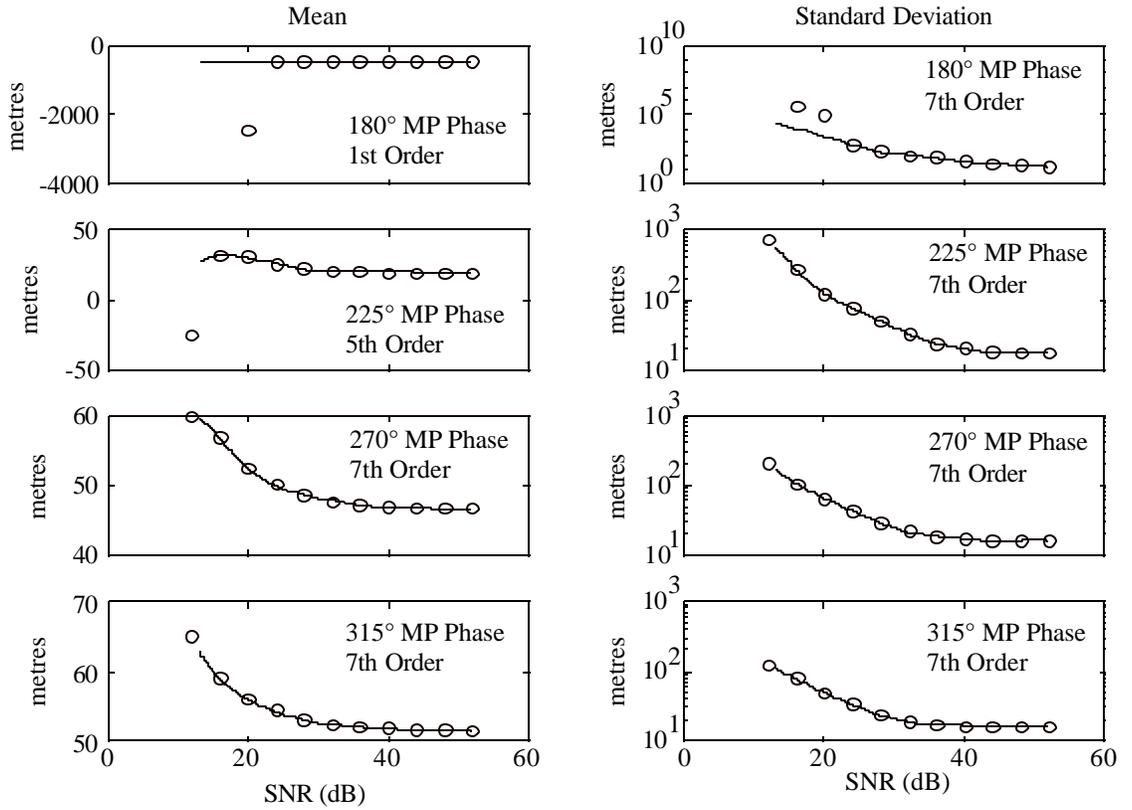
Figure 5.10b   MUSIC TOA Estimation Error Models (Urban)

('o' simulation points, '-' polynomial model)
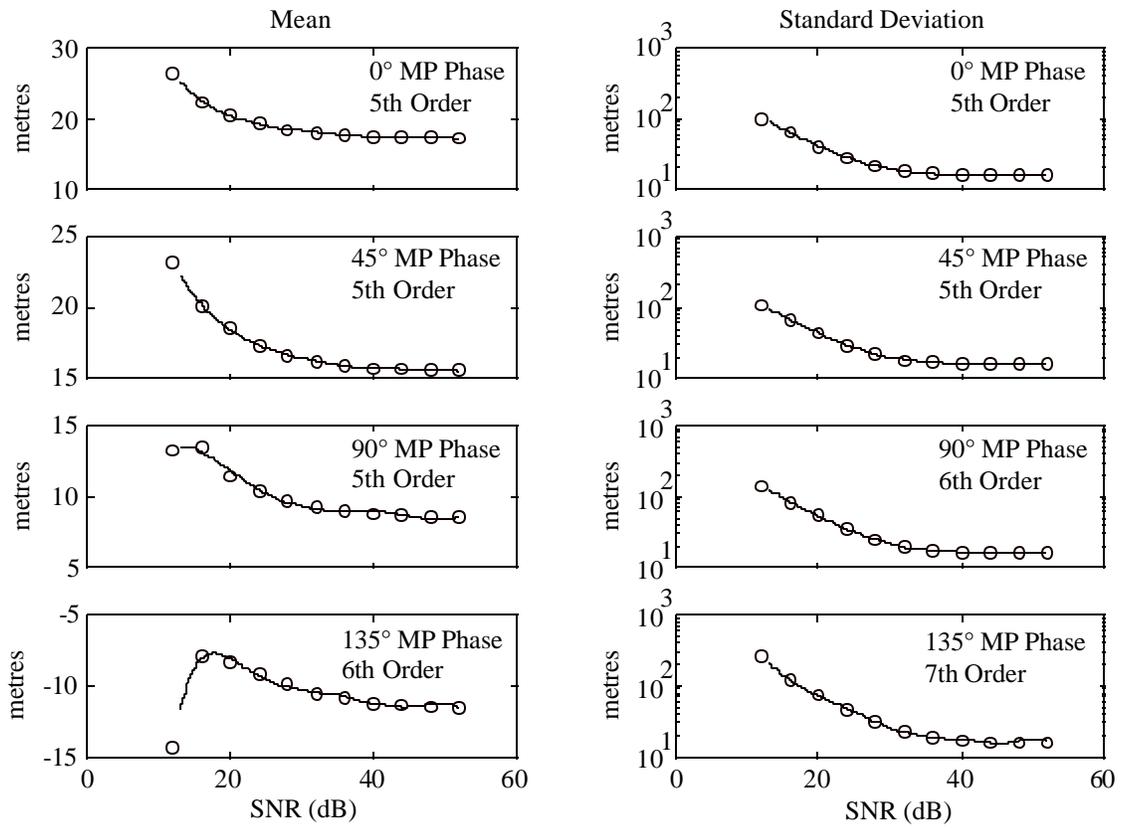
Figure 5.11a MUSIC TOA Estimation Error Models (Suburban)

('o' simulation points, '-' polynomial model)
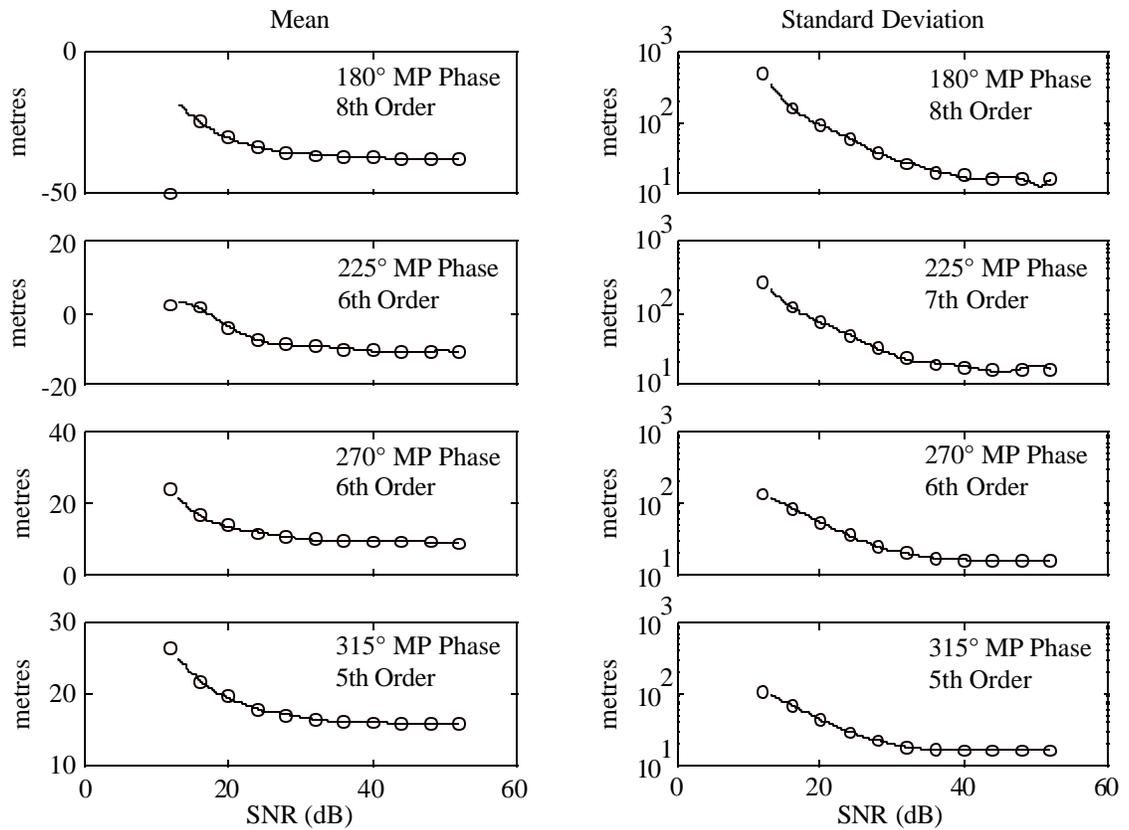
Figure 5.11b   MUSIC TOA Estimation Error Models (Suburban)

('o' simulation points, '-' polynomial model)

Once values for the TOA estimation error mean and standard deviation were randomly generated for a particular range, they were used to generate a TOA estimation error from the Normal distribution. The assumption of the normality of the TOA estimation errors was tested by generating normal probability plots for both the urban and suburban environments at various SNRs and for various multipath phases. Matlab® contains a function called normplot which allows one to asses whether a data set comes from a normal distribution. The normal probability plot for a suburban environment with multipath phase of 90º and a SNR of 15 dB is shown in Figure 5.12.
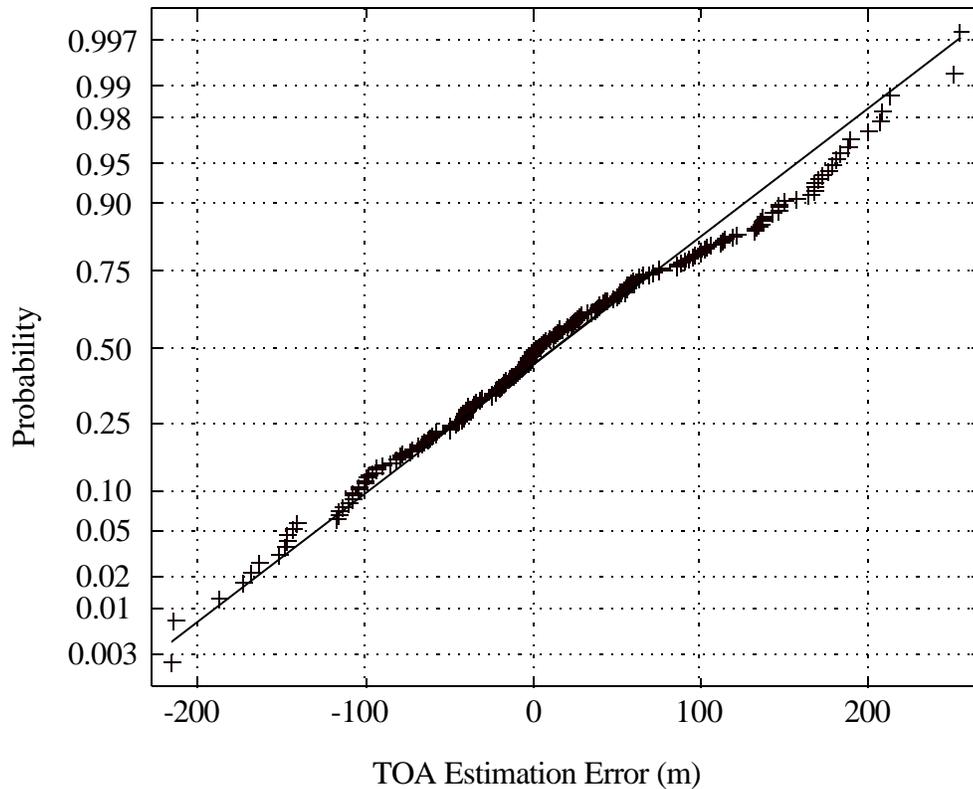
Figure 5.12   Normal Probability Plot for TOA Estimation Errors

Suburban Environment, 90° Multipath Phase, 15 dB SNR

        The crosses show the empirical probability for each TOA estimation error. The straight line connects the 25th and 75th percentiles of the data. Note that the tick marks on the y-axis are not uniformly spaced. Instead, the distances between them match the distances between the quantiles of a normal distribution. If the data samples fall near the solid line, it is reasonable to conclude that the data come from a normal distribution. For the data plotted in Figure 5.12, it is safe to assume that the TOA estimation errors are in fact normally distributed. Figure 5.12 is typical of the normal probability plots generated for other environments, phases, and SNRs. Therefore, it was considered safe to generate TOA estimation errors for the positioning simulations using the normal distribution.

CHAPTER 6

POSITIONING SIMULATIONS

6.1 Introduction

These simulations are based on the Telus Mobility cellular network in Calgary, Alberta, Canada. Only four Cellocate™ receivers were available at the time of this work. It was, therefore, impossible to install Cellocate™ receivers throughout the network and field test the system on a city-wide scale. This made it necessary to simulate system performance on a large scale.

This chapter describes the positioning simulations and the models used. The results of these simulations indicate the effect of various error sources as well as network geometry on position accuracy. The chapter proceeds by describing the Telus Mobility network, the multipath models used to corrupt ranges, the propagation loss equations used to determine SNR and the positioning algorithms. Following this, the results of the simulations are presented and analyzed.

6.2 Description of Positioning Simulations

6.2.1 Cellular Network

The Telus Mobility cellular telephone network was used to simulate the performance of the Cellocate™ system on a city-wide scale. When the simulation software was designed and written, the network consisted of approximately 40 cell sites. Almost all of the cell sites are sectorized in that they use directional antennas. To make the simulation computations reasonable, however, all cell sites were assumed to be isotropic. In addition, a small number of the cell sites share the same geographical location. For instance, three separate cell sites, with directional antennas pointing in different directions, are located on top of the Petro Canada building. These three cell sites were regarded as one isotropic site. Therefore, for the simulations, a total of 36 isotropic cell sites were assumed.

The relative positions of the 36 cell sites, with respect to an outline of the Calgary city limits, are given in Figure 6.1. The cell sites are represented by the small circles. Also

included in Figure 6.1 are two larger circles centered on downtown Calgary. The larger is of 12 km radius whereas the smaller is of 2 km radius. A grid of hypothetical cellular telephone locations was generated over the city area. The spacing between the grid points was 250 m. Since the extreme northern and southern portions of the cities are undeveloped and cell sites outside of the city limits were in general not used, a boundary of 12 km from the city centre was chosen. Only grid points within the 12 km radius were used in the simulations.

For the simulations of this chapter and the field tests of Chapters 7 and 8, all grid points and cell sites are assumed to lie in a horizontal plane. The position to be estimated is a two-dimensional, horizontal position. The heights of the grid points and cell sites are not accounted for.
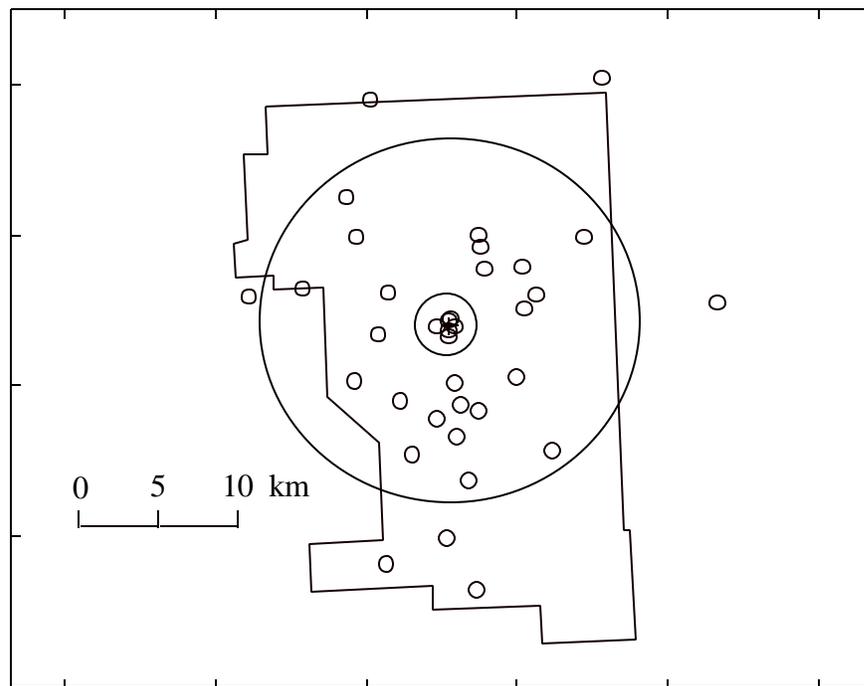
Figure 6.1   Telus Mobility Cellular Network in Calgary, Alberta, Canada

6.2.2 Propagation Environment

The 2 km radius circle serves as a boundary between propagation environments. Calgary consists of two main propagation environments. The downtown or city centre is densely developed with high-rise office buildings in excess of 20 stories. Towards the periphery of the city core the building size is somewhat smaller with the exception of some multi-story apartment buildings. The city core extends to a radius of approximately 2 km. Outside the city core the development is mostly suburban with the exception of the occasional multi-story building, such as on the university campus or in industrial parks. The city was therefore divided into the two distinct propagation areas - that within 2 km of the city centre is classified as urban whereas from 2 km to 12 km the classification is suburban. These two areas are distinct with regards to multipath and propagation loss. Therefore, different models will be used for these areas in the simulations.

Some grid points lying near the interface of the two propagation environments will of course be visible from cell sites in both areas. However, for any particular grid point, the same multipath and propagation loss models will be used for all cell sites, regardless of whether they reside in the urban zone or suburban zone. The models used depend on the location of the grid point. For grid points within the urban zone, this may make the simulation results somewhat pessimistic. However, the opposite effect will occur in the suburban zone. The overall effect of this simplification will, therefore, be insignificant.

Of the grid of hypothetical cellular telephone locations generated, 7211 fall within the 12 km radius of the city centre. Of those 7211 points, 195 fall within the 2 km radius whereas the remaining 7016 lie between the 2 km and 12 km radii. Each simulation run consists of one trial at each of the 7211 points. The result of each trial is a two dimensional error in the estimated position of a cellular telephone at that location. Statistics for the simulation run are then generated from the two dimensional error at each of the grid points.

6.2.3 Simulated Multipath

The simulations proceeded by stepping through each of the 7211 grid points. For each point the horizontal LOS distance to each of the 36 cell sites was calculated. An excess range was then added to each LOS range in order to simulate multipath. The excess range

was generated from probability density functions (PDFs) based on Turin's experiments discussed in section 2.3 (Turn et al., 1972a). In the experiments, the van containing the receiver was driven through four different propagation environments. Of interest here are Turin's areas A and B. Area A consisted of narrow streets and densely packed buildings up to 50 stories high. Turin describes it as "typical of the worst propagation environment one could expect in a modern metropolis". Note that some 25 years later, buildings well in excess of 50 stories are common in urban centres. Within this area multipath and diffraction are common due to the abundance of reflecting objects. Propagation loss is also high. This area was chosen for the 'urban' area of Calgary. Turin's Area B was downtown Oakland California. This area consisted of "sparsely clustered skyscrapers, up to 40 stories, interspersed with 2-3 story ... buildings." He notes that LOS transmission was possible if not common in this area. Multipath is, however, also common. This area was chosen for the 'suburban' area of Calgary - that region from 2 km radius to 12 km radius. Although this choice is somewhat pessimistic, the 'suburban' part of Calgary consists of true suburbia, as well as industrial parks, apartment buildings and so on. Turin's area B is, therefore, a reasonable choice.

Morley (1995) created PDFs from Turin's data. Turin plots probability-of-occupancy curves based on his empirical data. These plots (Fig. 4 in Turin et al. 1972a) show the empirical probability of a path being within ±50 ft of an excess range between 0 ft and 7000 ft. The results are divided into bins of width 100 ft. Morley derived the PDFs of the first path excess range from these curves. To determine the probability that the first path arrives in some bin, he multiplied that bin's probability of occupancy by the probability of non-occupancy of all the preceding bins. From these functions, the probability that the first signal to arrive at the receive antenna will travel a particular excess distance due to reflection or refraction, can be determined.

Plots of Morley's PDFs for Turin's areas A and B are shown by the dashed curves in Figures 6.2 and 6.3 respectively. These figures were traced from Morley's Figures 2-3 (a) and (b). Also plotted in these figures are curves generated from 10th order polynomials using the Matlab® polyval function. The polynomial coefficients were determined from

points on Morley's curves with the Matlab® polyfit function. The resolution of the polynomial curves in Figures 6.2 and 6.3 is 1 ft. Values of the polynomials in these two figures are occasionally negative. This is not practically possible since these curves represent probabilities. This is later accounted for in the method of generating random numbers from these polynomials.

The PDF used to generate an excess range depended on whether the grid point under consideration was located in the urban zone or the suburban zone. Once the appropriate PDF was chosen, it was used to generate a random excess range according to that distribution. The rejection method (Matlab®, 1994a) was chosen to generate a random number according to one of the two PDFs. Let F be the distribution with corresponding PDF $f(x)$ from Figure 6.2 or Figure 6.3. The reduction method requires another distribution G with PDF $g(x)$ and a constant c such that

$$f(x) \leq cg(x) \quad \forall x. \tag{6.1}$$

Random numbers from the distribution F may then be generated by the following steps:

1. Generate a random number x from distribution G.

2. Calculate $r = \dfrac{cg(x)}{f(x)}$.

3. Generate a uniform random number u.

4. If $ur < 1$ then x is a random number from distribution F.

5. If $ur \geq 1$ then repeat steps 1 to 3.
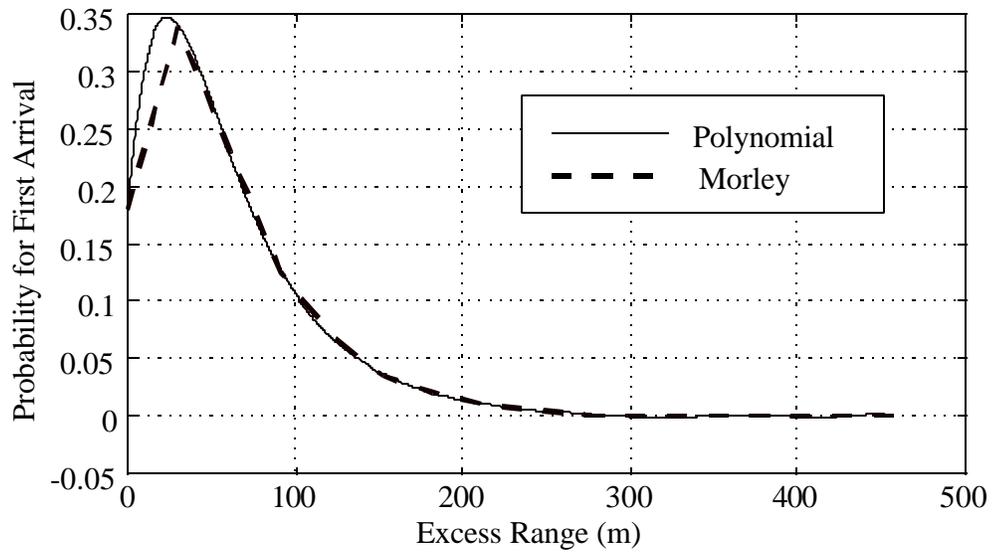
Figure 6.2   Probability Density Function of First Arrival (Urban)
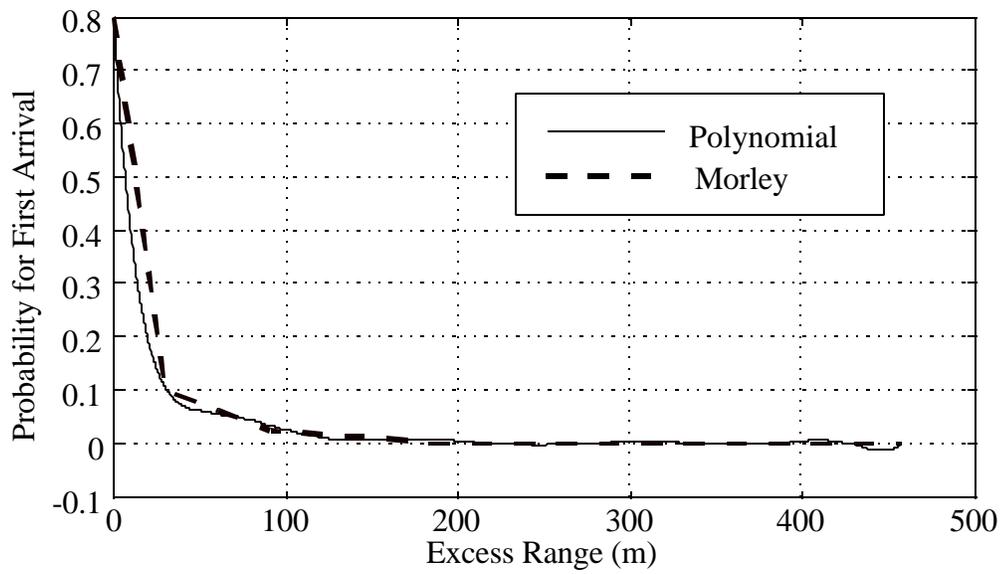
Figure 6.3   Probability Density Function of First Arrival (Suburban)

The uniform distribution was chosen for G and a value of 1/1500 was given to $g(x)$. The constant c was set to 550. A value of 1/1500 was chosen for $g(x)$ since Turin's probability-of-occupancy curves, and hence Morley's PDFs, used a domain of [0, 1500 ft]. The generation of random excess ranges was done with the same domain and the result converted from feet to metres. The first step then was to generate a uniformly distributed number x between 0 and 1500. The value of $f(x)$ was then evaluated using the appropriate polynomial model. As previously noted, the polynomial models shown in Figures 6.2 and 6.3 are occasionally negative. If for x the polynomial value was negative, $f(x)$ was set to a very small positive number (1e-10). A uniform random number of range [0,1] was then generated and steps 4 and 5 of the method followed.

How well this method generates random variables according to Morley's PDFs is illustrated in the histograms of Figure 6.4 and Figure 6.5. Shown in each of these figures are Morley's PDF and a histogram of 100,000 random numbers generated with the reduction method. The bin width in all cases is approximately 30.5 metres which corresponds to the 100 ft bin width of the Turin data from which Morley derived his PDFs. Agreement between the Morley's PDFs and the histograms is good with the exception of the first bin. It is important to note that the reduction method histogram curves of Figures 6.4 and 6.5 characterize the multipath in the simulations to follow.
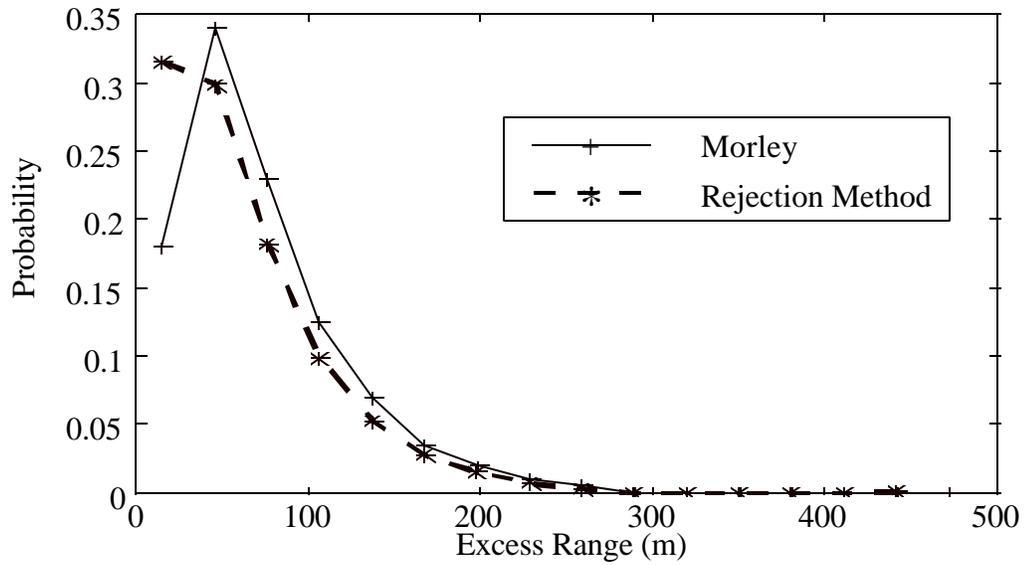
Figure 6.4   Excess Range Histograms (Urban)
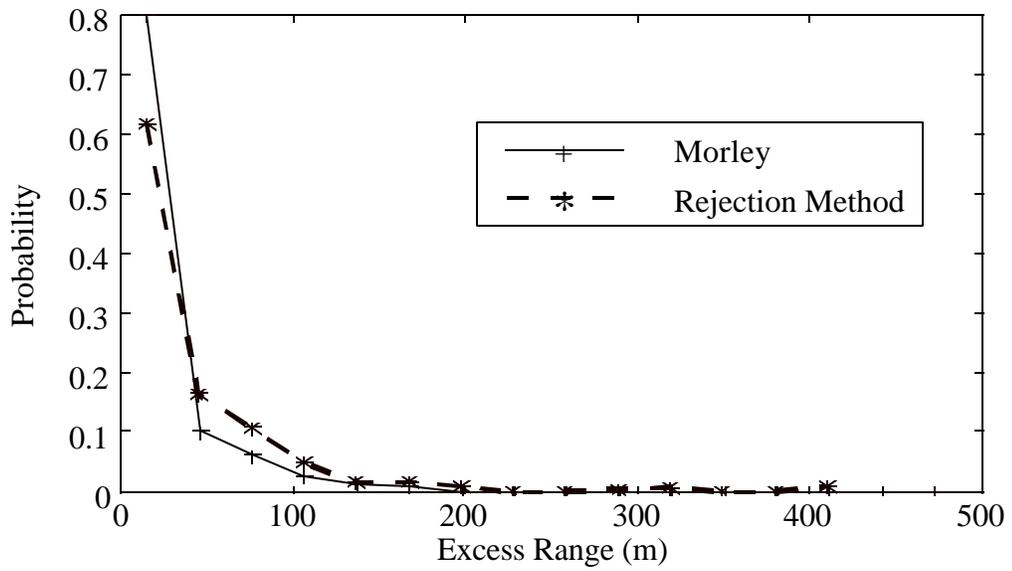
Figure 6.5   Excess Range Histograms (Suburban)

6.2.4 SNR Calculation

Once random excess ranges were generated and added to the LOS distance to each cell site, the propagation losses for the resulting ranges were calculated using Hata's empirical formulae (Hata, 1980). Hata's equations are based on empirical data and predict the propagation loss for land mobile radio services. The equations are considered applicable for the conditions in Table 6.1.

Table 6.1   Conditions for Hata's Propagation Loss Equations

Frequency	150 MHz - 1500 MHz
Base Station Antenna Height	30 m - 200 m
Vehicular Station Antenna Height	1 m - 10 m
Distance	1 km - 20 km

Hata presents a standard formula for urban propagation loss with corrections for vehicle station antenna height and other propagation environments. The standard formula is

$$L_{urban} = 69.55 + 26.16 \cdot \log_{10} f_c - 13.82 \cdot \log_{10} h_b - a(h_m)$$
$$+ \left(44.9 - 6.55 \cdot \log_{10} h_b\right) \cdot \log_{10} R$$

(6.2)

where L = propagation loss in dB,

f_c = carrier frequency in MHz,

h_b = base station antenna height in m,

h_m = vehicular station antenna height in m,

$a(h_m)$ = correction for vehicular station antenna height in dB,

and

R = propagation distance in km (LOS + excess range).

Two different vehicular antenna height corrections are given. For a 'medium-small' city the correction is

$$a(h_m) = (1.1 \cdot \log_{10} f_c - 0.7) \cdot h_m - (1.56 \cdot \log_{10} f_c - 0.8) \quad (6.3)$$

whereas for a large city the correction is

$$a(h_m) = 3.2 \cdot (\log_{10} 11.75 h_m)^2 - 4.97, \qquad f_c \geq 400 \text{ MHz.} \quad (6.4)$$

The standard equation (6.2) is corrected for propagation loss in a suburban area. The modified equation is

$$L_{suburban} = L_{urban} - 2\{\log_{10}(f_c/28)\}^2 - 5.4. \qquad (6.5)$$

For the urban area of the positioning simulations (radius up to 2 km), equations (6.2) and (6.4) were used. For the suburban area (2 km to 12 km radius) equations (6.3) and (6.5) were used. As shown in Table 5.2, the carrier frequency assumed was 840 MHz. The cell site antenna height, h_b, was set to 50 m and the cellular telephone height, h_m, was set to 1.5 m.

Following Morley, an additional path loss due to the log-normal distribution of signal strength, was added to the loss calculated from Hata's equations. Turin reports that path strength as a function of excess range is log-normally distributed. From Turin's plots of path strengths, Morley determined that the standard deviation for the areas of interest here is approximately 6 dB. Therefore, additional path losses were generated from normally distributed random numbers with mean of 0 dB and standard deviation of 6 dB.

Short term fading was taken into account by incorporating a fade margin in the SNR calculation. A fade margin of 8 dB ensures that signal loss due to short term fading happens only 10% of the time (Lee, 1982). Long term fading was not accounted for. Antenna gains and cable losses were also incorporated after Morley. Therefore, the SNR in dB was calculated by

$$SNR = P_t - L + C - F - N \qquad (6.6)$$

where P_t = transmit power = 28 dBm (nominal class III power),

L = path loss in dB (Hata's equations + Turin's log-normal path strength),

C = mobile ant. gain + cell site ant. gain + cable losses,

= 2.2 dB + 6 dB - 5 dB,

F = fade margin = 8 dB, and

N = noise floor at 30 kHz = -129 dBm.

Obviously all 36 cell sites in the city would not be able to receive a signal from a cellular telephone anywhere within the city. An SNR threshold was set to determine which cell sites could receive the signal and be used in the positioning calculations. Hand-off information from Telus Mobility[1] for the Calgary area indicates that, on average, a cellular telephone is 'visible' from seven cell sites. To obtain an average number of approximately seven participating cell sites, a threshold of 13 dB was chosen. One measurement made by Telus Mobility in the vicinity of the cell sites used for the field tests of Chapter 8, yielded an SNR of 24 dB. A minimum SNR of 13 dB is, therefore, realistic. For the simulations to follow, a threshold of 13 dB resulted in a mean of 8.2 cell sites (and therefore observations) with standard deviation of 3.9.

6.2.5 Simulated TOA Estimation Error

Once the participating cell sites were determined, the range to each cell site, already corrupted by multipath, was further corrupted by the TOA estimation error of root MUSIC. The polynomial models of Figures 5.10 and 5.11 were used for this purpose. First, a uniformly distributed random phase was selected from the set {0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°}. The corresponding polynomial model, based on whether the grid point was in the urban zone or the suburban zone, was then used to determine the TOA estimation error mean and standard deviation for the given SNR. Since the polynomial models are only valid up to 50 dB SNR, any calculated SNR above 50 dB was fixed to 50 dB. This is reasonable since most of the MUSIC error mean and standard deviation curves are constant after 50 dB. In addition, in very few cases was the SNR greater that 50 dB. A histogram of

SNRs for one simulation run is plotted in Figure 6.6. It is clear that over 95% of the SNRs are below 50 dB.
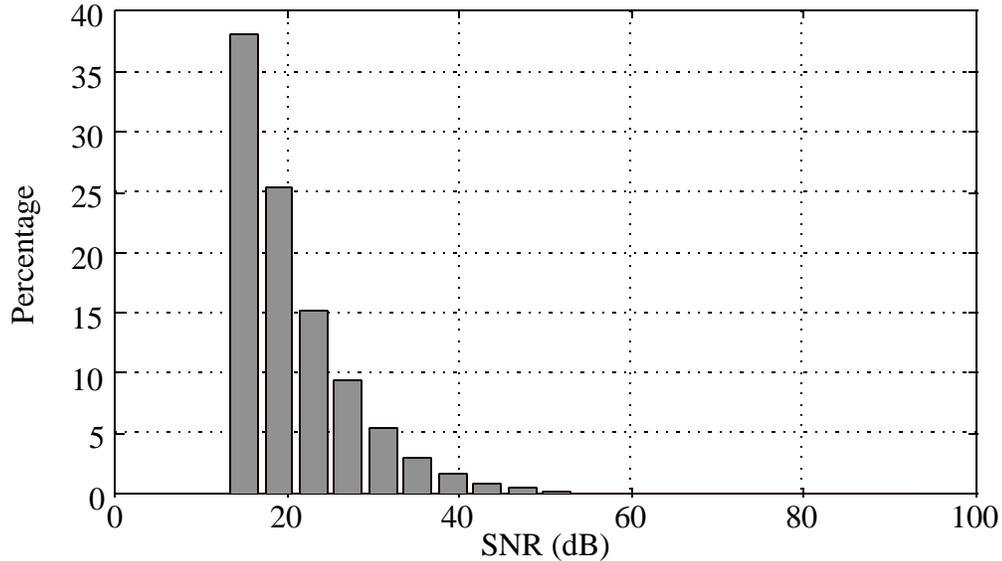
Figure 6.6   SNR Histogram for One Simulation Run

The TOA estimation error mean and standard deviation were used to generate a normally distributed random TOA estimation error. This error was added to the multipath corrupted range. The result for each grid point then was a set of corrupted ranges to those cell sites which were visible to that grid point according to the SNR threshold. These ranges were the observations used in the positioning estimation process.

6.2.6 Geometry

Given the positions of the grid points and cell sites involved in the position estimation process, the HDOP may be calculated from the design matrix and observation covariance matrix as discussed in Chapter 4. Due to the design of these matrices, circular and hyperbolic trilateration yield an identical HDOP value. The HDOP was calculated for every grid point within the 12 km radius. As discussed previously, the participating cell sites

---

[1] Private Communication with Brent MacArthur of Telus Mobility

for each grid point are a function of the multipath models used. Both the multipath models and propagation loss calculations require the generation of random numbers. Random number generators are based on a seed value, each seed generating a different series of random numbers. A simulation run, defined as one trial at each of the 7211 grid points, is based on a specific seed value for the normal and uniform random number generators. In Matlab®, the normal and uniform random number generators maintain separate seeds.

Therefore, for each simulation run, the corruption of the LOS ranges due to multipath and the propagation losses are different. The participating cell sites for a particular grid point may change from one run to the next. Therefore, the result of any one simulation run is only one realization of a random process. Figure 6.7 is a plot similar to that of Figure 6.1. Shown in Figure 6.7 are regions of similar HDOP for one particular realization. There is little difference in the HDOP values from one realization to the next and Figure 6.7 may be considered as typical. Within each region, the grid points have an HDOP within the approximate ranges indicated.

There are four regions. Within the 2 km radius the HDOP is relatively high. In this area there are a large number of cell sites in a very small geographical area. The urban propagation model will eliminate many of the cell sites outside of the urban zone from being used in the position fix due to high propagation loss. The result is poor geometry especially towards the fringe of the urban zone. The best geometry occurs primarily to the south and northeast of the city centre. The number of surrounding cell sites in these areas accounts for this. Throughout the rest of the city the HDOP is fairly uniform at about 2. The exceptions are the edges in the north, southeast, and southwest. The northern area corresponds to the location of the airport whereas the southern areas are essentially empty of development. Outside the 12 km radius, with the exception of due south, the land is undeveloped. Hence, the number of cell sites in those areas is small and the HDOP along the edge of the 12 km radius is high.

Region **D**
HDOP > 8

Region **B**
1 < HDOP < 3

Region **A**
0 < HDOP < 1

Region **C**
HDOP > 5

0   5   10 km
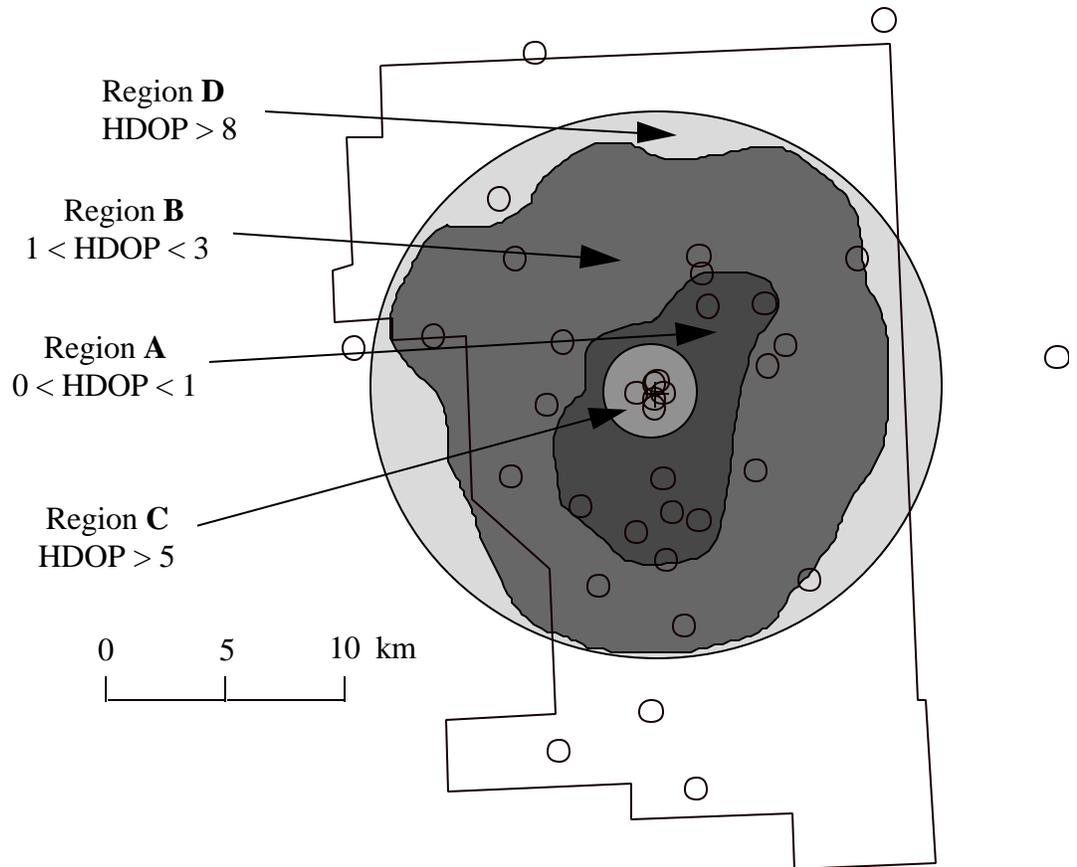
Figure 6.7   HDOP Regions for One Particular Simulation Run

A histogram of the HDOP values appears in Figure 6.8. It is clear that over 50% of the grid points have an HDOP of 1 or less and 80% have an HDOP of 2 or less. A total of 121 grid points had an HDOP greater than 10, the largest being 221. From Figures 6.7 and 6.8, one concludes that the geometry for the vast majority of the test area is excellent. It is emphasized, however, that these results are based on an SNR threshold of 13 dB and the propagation loss equations (6.2) through (6.5).
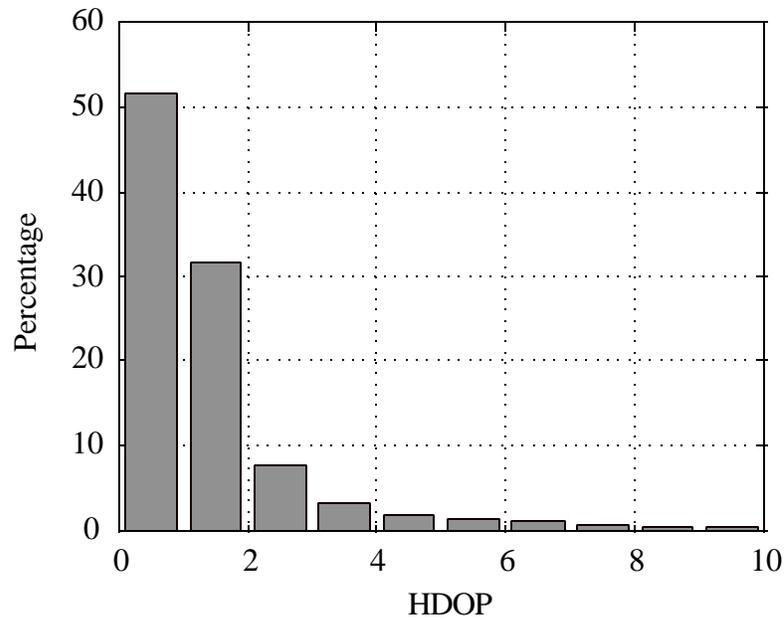
Figure 6.8   HDOP Histogram for One Simulation Run

6.2.7 Position Estimation

6.2.7.1 Initial Position Estimate

Iterative methods such as LS require an initial position to begin the iterative process. For many applications, due to geometry and the distances involved, the accuracy of the initial solution need not be very good. For example, in the case of GPS the transmitting sources are approximately 20,000 km from the receiver and the satellites will be within 180° of each other in the celestial sphere. Errors in the initial position on the order of the radius of the earth may, therefore, be tolerated. For cellular positioning, however, the propagation distances are in general less than 5 km and the receiver is often located amongst the transmitters. This results in the possibility of multiple or ambiguous solutions within close proximity of one another. This situation requires a much more accurate initial position.

Consider the case of Figure 6.9. This figure shows a simulation grid point and the only three cell sites which exceed the 13 dB threshold. The grid point is indicated by the

asterisk, the cell sites are marked by filled circles and the city limit is drawn by the dashed line. Also shown are the hyperbolic lines of position corresponding to the range differences. In this case the ranges of cell sites A and C were differenced with that of cell site B. The resulting hyperbolas are identified as AB and CB. The ranges from the grid point to the cell sites were uncorrupted. As a result, the two hyperbolas intersect at the true grid point position. Of interest here is the fact that the two hyperbolas intersect at two points within 2 km of one another. In many applications the second solution is so far remote that it can be easily discounted. Figure 6.9 shows, however, that for the geometry and distances involved in this application, it is possible to have two reasonable solutions.
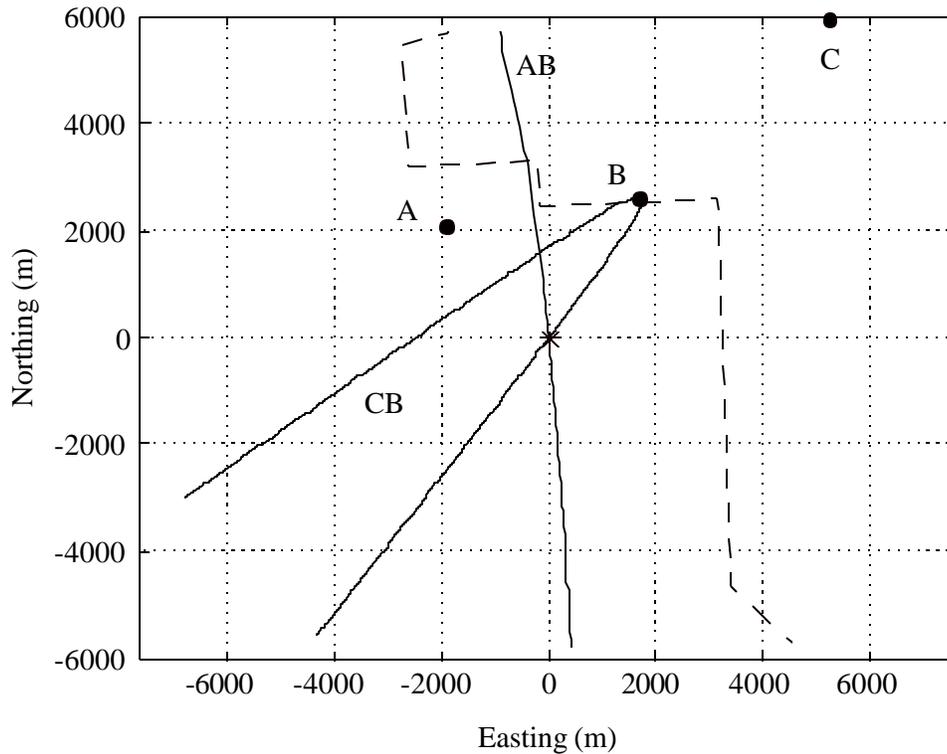
Figure 6.9   Dual Solutions for Two Intersecting Hyperbolas
HDOP = 15.5, NDOP = 15.4, EDOP = 1.2

(range B differenced from ranges A and C)

To which solution LS converges is a function of the initial position provided to LS. Figure 6.10 is a replica of Figure 6.9 with some additional information. Shown in Figure 6.10 is a grid of 289 initial positions, spaced by 500 m. These initial positions were used, one after the other, to initialize the LS iterative process. This resulted in 289 LS solutions. The result of each solution is indicated by the symbol used to mark the initial position used for that particular case. Three categories of solutions were possible. LS was found to either converge to one of the two solutions or to diverge. Those initial positions which resulted in convergence to the correct solution are indicated by crosses; those that resulted in convergence to the second solution are marked by dots; those resulting in divergence are indicated by circles. The correct solution is again indicated by an asterisk whereas the second solution is marked by an x.
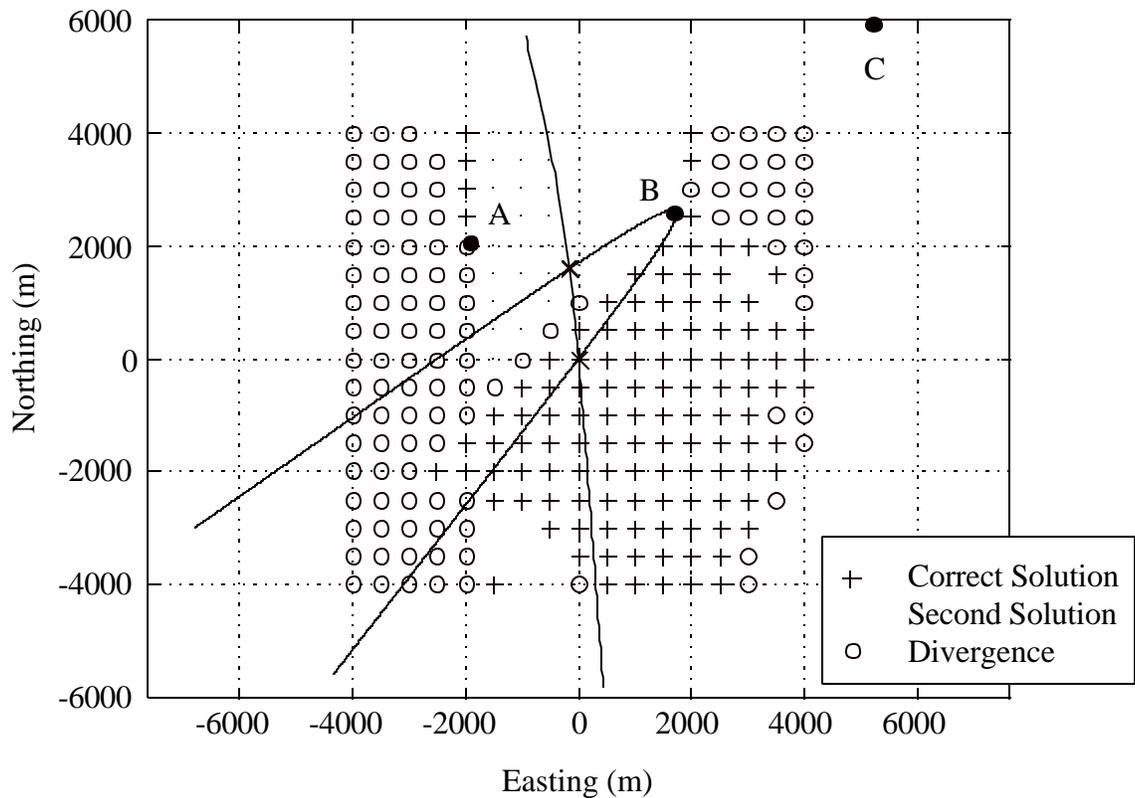
Figure 6.10   Result of Various Initial Positions in the Case of Two Solutions
(range B differenced from ranges A and C)

Of the 289 initial positions, 40% resulted in convergence to the correct solution, 22% resulted in convergence to the second solution and the remaining 38% resulted in divergence. It is intuitively obvious that those initial positions close to one solution or the other will cause convergence to that solution. This is confirmed by the quiver plot of Figure 6.11. This plot illustrates vectors corresponding to the corrections to the position parameters after one iteration. Each arrow represents the parameter correction given the tail of the arrow as the initial position. From this plot one can easily see that the solution eventually

112

arrived at depends on the initial position. Since the LS process is the minimization of some objective function, one may add a third dimension to Figure 6.10 representing the value of the objective function. The three dimensional surface so obtained will have two minima - one at each of the intersections. Which of the two minima LS 'slips into' depends on where LS starts.
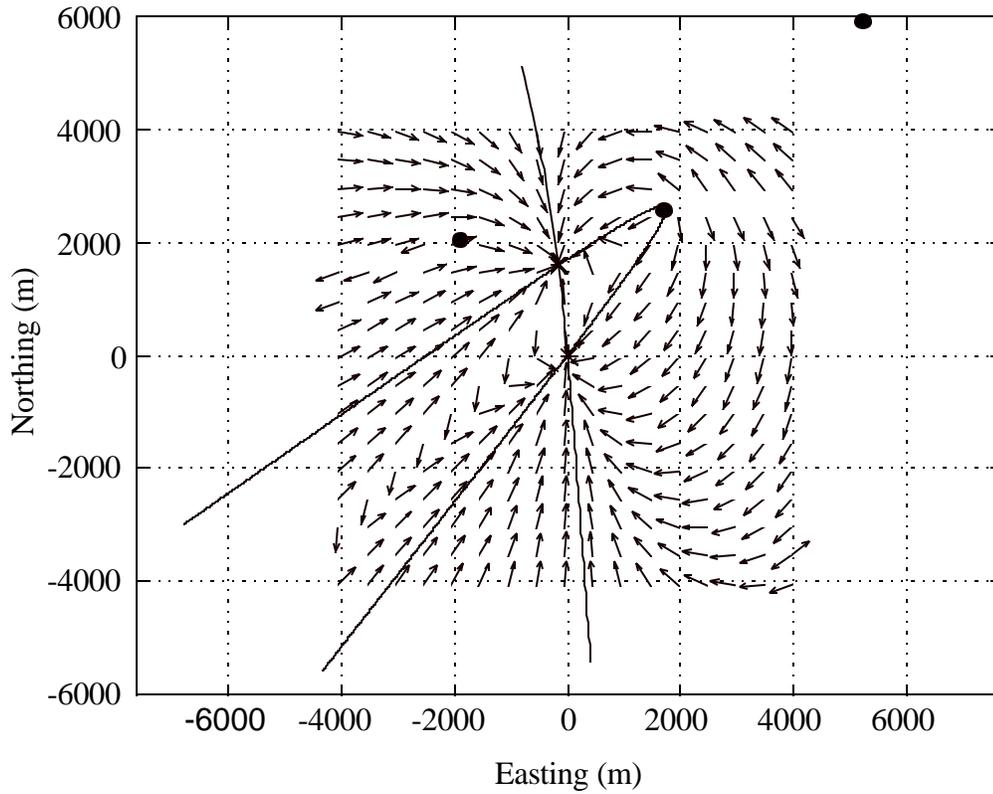
Figure 6.11   Parameter Correction Vectors

In contrast to Figure 6.10, Figure 6.11 gives the impression that an overwhelming majority of the initial positions result in convergence to one solution or the other. This is due to the vectors of Figure 6.11 being of unit magnitude. In reality, the magnitudes of those vectors corresponding to the circles of Figure 6.10 are orders of magnitude greater than those

corresponding to the dots and crosses. In other words, although many of the vectors point toward one of the solutions, their magnitude is so large that they greatly overshoot the solution resulting in divergence.

The cell site used as reference for differencing does not affect the outcomes illustrated in Figures 6.9 through 6.11. Figure 6.12 shows the results when the ranges of cell sites B and C are differenced from that of A. The results are identical to those of Figure 6.10. In this case the two hyperbolas intersect at a very shallow angle but at the
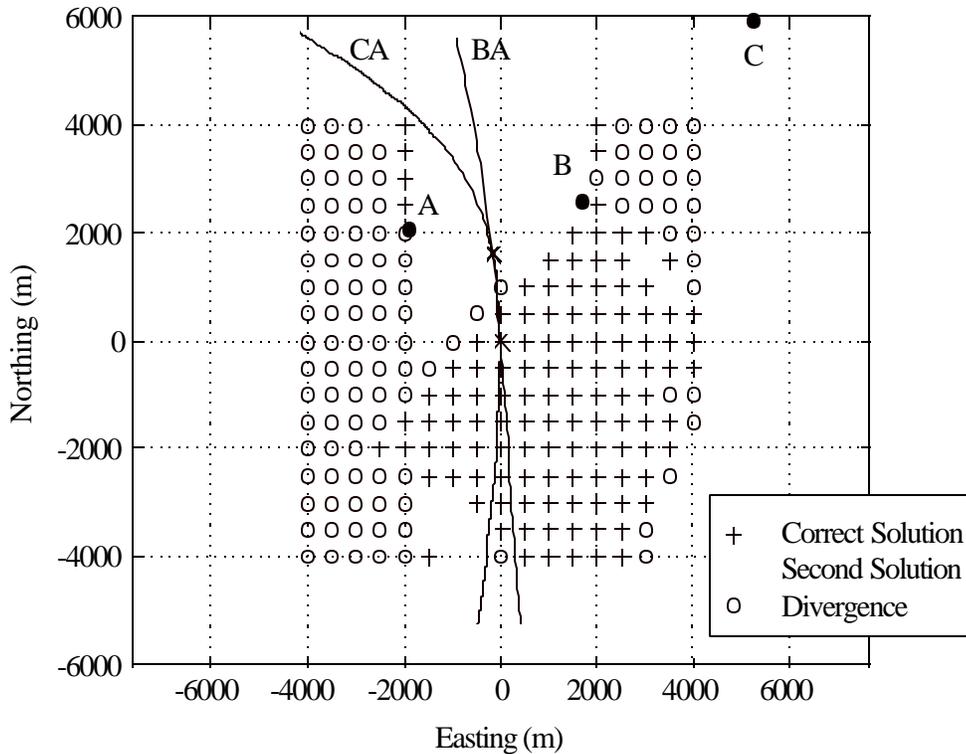
Figure 6.12   Result of Various Initial Positions in the Case of Two Solutions
(range A differenced from ranges B and C)

same two points. It is obvious that the hyperbola corresponding to range difference AC will also intersect that of range difference BC at the same two points.

The preceding discussion has demonstrated the possibility that in the case of poor geometry and three TOAs, two solutions in close proximity of one another are possible. The solution arrived at by LS is largely a function of the initial position. One would expect that more observations will eliminate this problem. In general this is true; although possible, it is highly improbable that n hyperbolas formed from $n+1$ TOAs (where $n > 2$), will intersect at two points. However, it is possible that some subset of hyperbolas will approximately intersect at a second point and that LS may converge to this point. Such a case is illustrated in Figure 6.13. This is another example from the simulation run which provided the previous example.
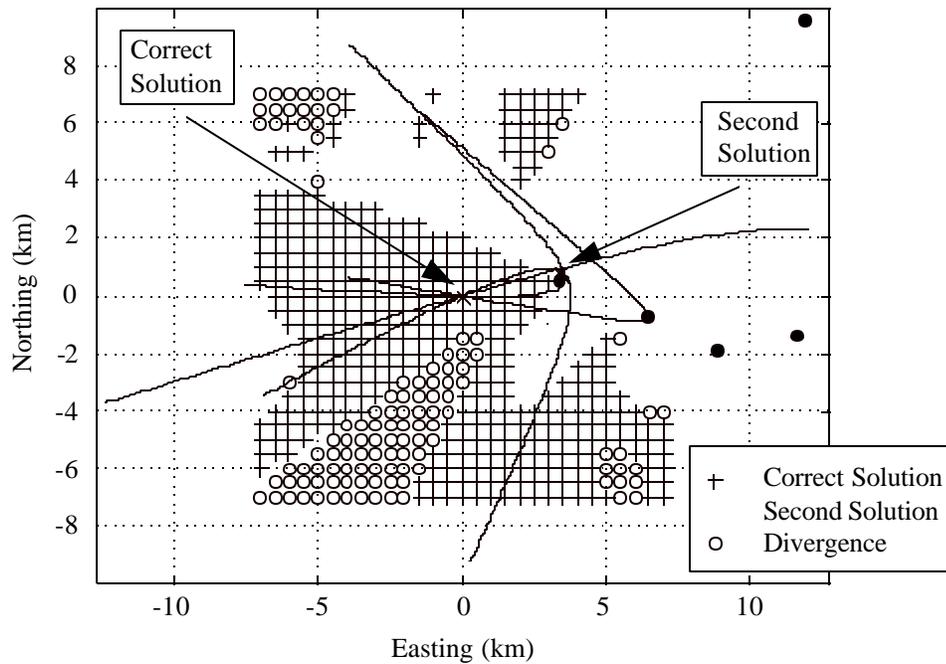
Figure 6.13   Result of Various Initial Positions in the Case of Four
                Intersecting Hyperbolas
                HDOP = 13.2, NDOP = 3.5, EDOP = 12.7

In Figure 6.13, the four hyperbolas are formed by differencing between the five TOAs in a sequential manner rather than differencing four of the TOAs from the remainder. Note that the HDOP is very high indicating poor geometry. The correct solution is at coordinates (0, 0) where all four hyperbolas intersect exactly. Three of the hyperbolas come very close to intersecting again at the point (3, 1). LS converges to this second solution for 41% of the initial positions shown. Forty-six percent of the initial positions resulted in convergence to the correct solution and the remaining 13% resulted in divergence. This demonstrates that even when redundancy exists, an accurate initial position is critical when geometry is poor.

Geometry in the form of DOP values, however, is not a reliable indicator for warning when multiple solutions may exist. Other examples, similar to those just presented, were found having HDOP values of approximately 2. The problem of solution bifurcation in GPS has been studied by Chaffee and Abel (1993). They present examples in which two GPS solutions become possible as a GPS receiver approaches a satellite. Unfortunately the DOP values are low and do not warn that a secondary solution is possible. In the context of developing a direct GPS solution for the case of four observations, Chaffee and Abel consider the notion of solution uniqueness. They developed a simple geometric test for solution uniqueness.

The work of Chaffee and Abel may be transferred to the cellular positioning case. Consider three cell sites with horizontal positions $\mathbf{C}_1$, $\mathbf{C}_2$, and $\mathbf{C}_3$ where $\mathbf{C}_i = \begin{bmatrix} x_i & y_i \end{bmatrix}$. The TOA estimated at the ith cell site is T_i. The origin of the coordinate system is first moved to the position of the first cell site. TOA differences are then formed using the first cell site as the reference. We then have

$$
\begin{aligned}
\Delta T_{21} &= T_2 - T_1 \text{ and} \\
\Delta \mathrm{T}_{31} &= T_3 - T_1.
\end{aligned}
\tag{6.7}
$$

Now let the coordinates of cell sites 2 and 3, for the coordinate system with origin at the first cell site, be

$$\mathbf{C}_2' = \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \end{bmatrix} \text{ and}$$
$$\mathbf{C}_3' = \begin{bmatrix} x_3 - x_1 & y_3 - y_1 \end{bmatrix}. \tag{6.8}$$

A matrix $\mathbf{Z}$ is formed where

$$\mathbf{Z} = \begin{bmatrix} \mathbf{C}_2' & \Delta T_{21} \\ \mathbf{C}_3' & \Delta T_{31} \end{bmatrix}. \tag{6.9}$$

Given that $\mathbf{Z}$ is of rank 2, the null space of $\mathbf{Z}$ will be one-dimensional. Let the unit vector spanning the null space of $\mathbf{Z}$ be $\mathbf{a}$. Chaffee and Abel show that, assuming a direct solution exists, there will be a unique solution to the hyperbolic equations when

$$\langle \mathbf{a}, \mathbf{a} \rangle < 0 \tag{6.10}$$

and two solutions exist when

$$\langle \mathbf{a}, \mathbf{a} \rangle > 0. \tag{6.11}$$

The operation $\langle \mathbf{a}, \mathbf{a} \rangle$ is the Lorentz inner product defined by,

$$\langle \mathbf{a}, \mathbf{a} \rangle = \left(\sum_{i=1}^{2} a_i^2 \right) - a_3^2. \tag{6.12}$$

Chaffee and Abel refer to $\langle \mathbf{a}, \mathbf{a} \rangle$ as the bifurcation parameter. Bifurcation of the solution occurs as the parameter passes through zero from positive to negative.

Note that the bifurcation parameter test only applies to an exactly determined system of equations. For horizontal cellular positioning this corresponds to two TOA

117

differences or two hyperbola, as in Figure 6.9. For the case of Figure 6.9, the bifurcation parameter was 0.45 indicating that two solutions exist. When redundancy exists, the bifurcation parameter must be calculated for each hyperbola pair.

As an example consider Figure 6.14. Cell site 3 is used as the reference such that hyperbola H13 is formed from the TOA differences between sites 1 and 3, H23 from the TOA differences between sites 2 and 3 and so on. Note that hyperbola H23 and H43 intersect at two separate points, the correct solution and the second solution. The bifurcation parameter for these two hyperbola is 0.05. Hyperbola H13 and H23 intersect only once as do H13 and H43. The bifurcation parameters for these hyperbola pairs are -0.005 and -0.08 respectively. When only hyperbola H23 and H43 are used, LS converges to either the correct solution or the second solution depending on the initial position. When all three hyperbola are used either the correct solution or Solution A are obtained, again depending on the initial position.
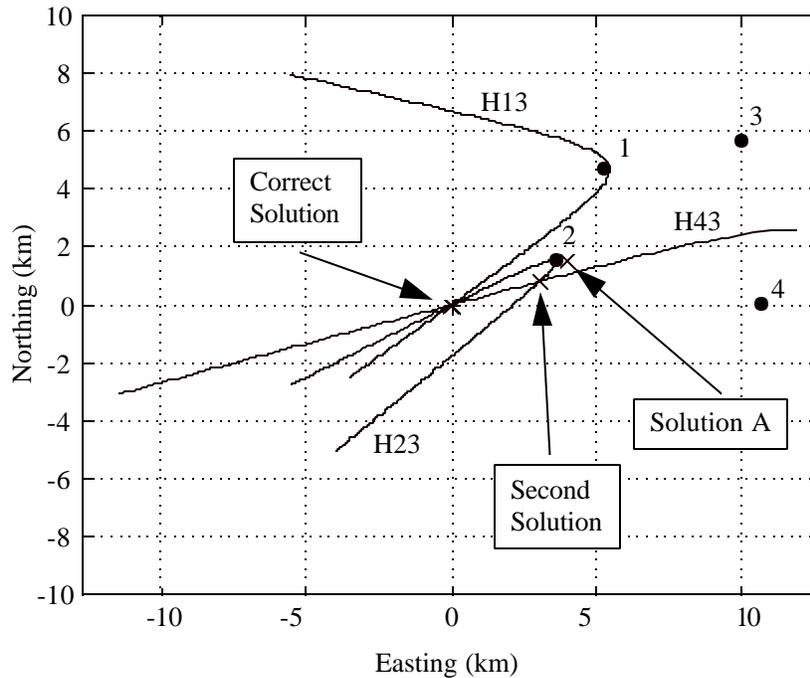
Figure 6.14   Multiple Solutions for Three Hyperbolas

$$\text{HDOP} = 17, \ \text{NDOP} = 6, \ \text{EDOP} = 16$$

Detection of the presence of two solutions is the obvious first step in dealing with this phenomena. The next problem is how to avoid convergence to a secondary solution. One possibility is to not make use of the offending hyperbola pair. However, since redundancy generally improves the position estimate, this is not a desirable course of action. A more attractive approach may be to assign less weight to the observations in question.

In any case, a very accurate initial position is desired. Due to the highly mobile nature of cellular telephony, past position is of little help in determining an initial position. Perhaps the easiest and most convenient source of an initial position is the cell site with which the cellular telephone communicates. The assumption could be made that the telephone is located within that cell and perhaps the geographical centre could be taken as the initial position. As cells become smaller due to increased load, this may result in a position of adequate accuracy given that the above assumption is correct. However, in many cases, the cell site with which the telephone communicates may not be the closest site. For example, in the field tests conducted, the telephone would often lock onto the cell site located atop the Petro Canada building, some 4 km in the distance, instead of a cell site only 1 km away. The reason was signal blockage. The Petro Canada building is a sky scraper in the central part of Calgary and the cell site antenna on its roof is visible from a large portion of the city. Although most of the other cell site antennas are mounted on masts, signal blockage and fading can occur. Therefore, it cannot be guaranteed that the cell site handling the telephone's call is the closest one.

A closed form algorithm, requiring only TOA observations, is the preferable solution. The LS plane intersection method discussed in Chapter 4 is investigated here for this purpose. A simulation run was executed using plane intersection to estimate position. Of the 7211 grid points, 6471 had four or more participating cell sites and their positions could therefore be estimated. The mean number of participating cell sites was eight with a standard deviation of 3.9.

Table 6.2 presents the results of the simulation run. The performance measures used here and in the simulation results to follow are Distance Root Mean Square (DRMS),

horizontal precision, mean horizontal error, and the 67th and 95th percentiles. Horizontal precision is defined as

$$\text{Hor. Prec. } = \sqrt{s_x^2 + s_y^2} \qquad (6.13)$$

where s_x is the standard deviation of the errors in the x component and s_y is the same for the y component errors. DRMS is an accuracy measure which is calculated by

$$\text{DRMS} = \sqrt{\frac{\sum_{i=1}^{N}\left[(\hat{x}_i - x_i)^2 + (\hat{y}_i - y_i)^2\right]}{N}} \qquad (6.14)$$

where $(\hat{x}_i, \hat{y}_i)$ is the position estimate of the ith grid point, (x_i, y_i) are the true coordinates of grid point i, and N is the number of grid points. The mean horizontal error is self-explanatory and the 67th percentile is that horizontal distance which 67% of the horizontal errors are less than. The values of the DRMS and 67th percentile suggest that there are a small number of large errors.

Table 6.2   Test Results for the LS Plane Intersection Method

DRMS (m)	2784
Horizontal Precision (m)	2784
Mean Hor. Error (m)	452
67th Percentile (m)	190
95th Percentile (m)	1237

Plotted in Figure 6.15 are the plane intersection DRMS, horizontal precision, and 67th and 95th percentiles for grid points with HDOP less than or equal to the abscissa. Also

plotted are the percentage of points, out of the 6471 total, and the mean number of participating cell sites for those points, with HDOP less than or equal to the abscissa. Recall that the results presented are for those grid points at which there were four or more TOA observations.

In general, the plots indicate the dependence of accuracy upon geometry. Performance is best when HDOP is small indicating good geometry. Geometry itself is a function of the number of participating cell sites as indicated by the plot of the mean number of cell sites versus HDOP. There is a significant jump in the DRMS and horizontal precision between HDOPs of 2 and 3 and an even larger jump between an HDOP of 5 and 6. These jumps may be correlated with the HDOP regions of Figure 6.7. The jump in DRMS between an HDOP of 2 and 3 corresponds to the difference between region B and region A in Figure 6.7. The jump in DRMS when HDOP increases from 5 to 6 may be associated with the difference between region C, the downtown area, and regions A and B. Those points with HDOP of 5 or less represent 93.5% of the total number of points whereas 95% have an HDOP of 6 or less. The increase in DRMS from approximately 1000 m to 2000 m is due to points consisting of only 1.5% of the total. A large majority of those points are in the downtown area which contains approximately 2.7% of the total number of points.

From these results, it appears that the LS plane intersection method is suitable for determination of initial position. The multiple solution cases presented in Figures 6.9 through 6.13 suggest that LS should be started as close as possible to the true solution to avoid convergence to a secondary solution. How accurate the initial position must be to ensure this, depends on the geometry of the particular case. In Figure 6.12, an accuracy of about 700 m, half the distance between the true and second solutions, is required. In the case of Figure 6.13, the required accuracy is 2 km, about two thirds of the distance between the two solutions.

The need to resolve between the two solutions depends on their proximity to one another. The two cases presented earlier are by no means exhaustive. As multiple solutions become closer to one another, it becomes less important to which solution LS converges. Convergence to the ambiguous solution, although undesirable, is not catastrophic if the two

solutions are within say 200 m of each other. However, in all cases, it would be most beneficial to provide some sort of reliability measure which could alert the user to the possibility of convergence to a secondary solution.

The LS plane intersection method was found to be numerically stable. For a simulation run in which uncorrupted ranges were used, the position estimated was accurate to within machine precision for the 6471 grid points which had four or more observations.
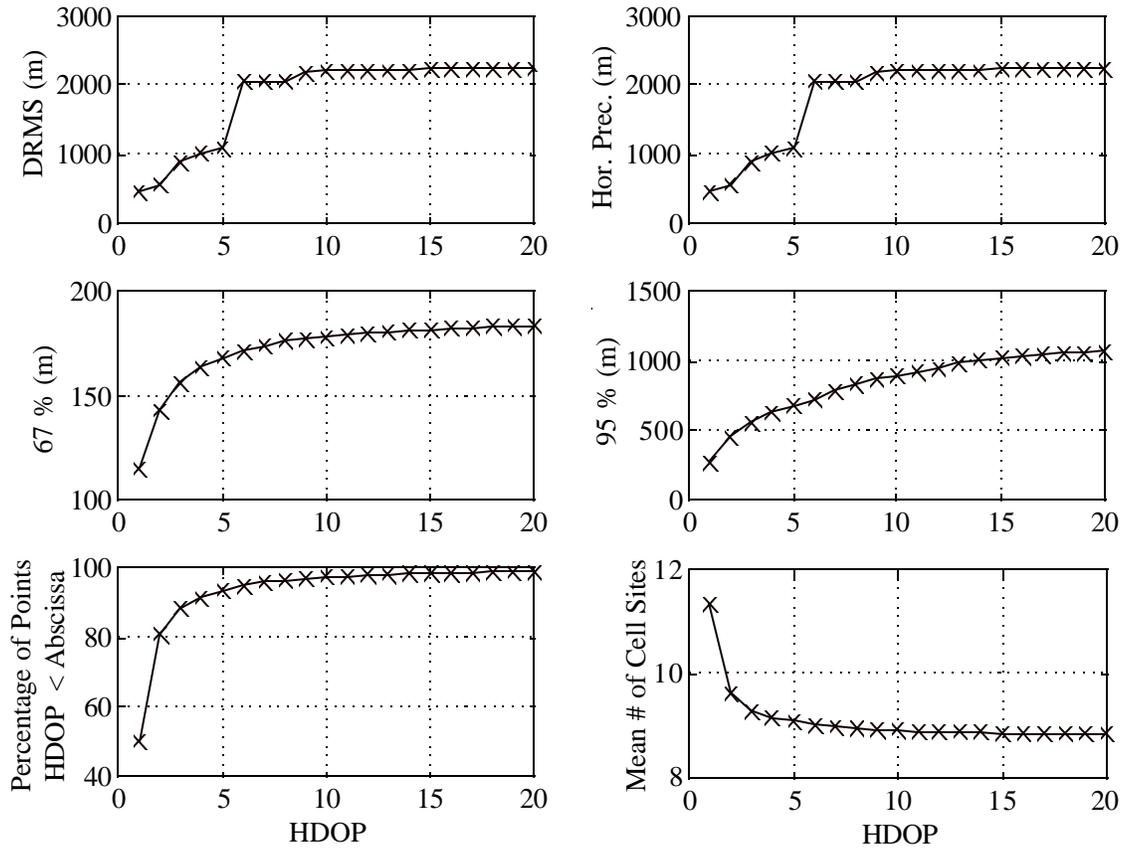
Figure 6.15   Performance Measures, Percentage of Points, and Mean Number
                                of Participating Cell Sites vs. HDOP for LS Plane
Intersection

Therefore, the initial position provided to LS in the simulations to follow was provided by the LS plane intersection method. If a grid point had four or more participating cell sites, its position was estimated with LS plane intersection. This position was then used to start the iterative process of LS. If LS diverged, or did not converge within 500 iterations, or the matrix $\left[\mathbf{A}^T \mathbf{C_l}^{-1} \mathbf{A} \right]$ became singular, the coordinates, plus 100 m in an arbitrary direction, of the participating cell site with the highest SNR, were used as the initial position. In the event that that this again resulted in no solution, that grid point was identified as such.

6.2.7.2 Least Squares Position Estimation

Four different LS algorithms were tested by simulation. Algorithms ls1 and ls3b employ circular trilateration - both horizontal position and TOT are solved for. They require an initial estimate of the TOT. Since the simulations are based on ranges and TOA was calculated by dividing range by the speed of light, the initial TOT estimate provided to LS was always zero. The two algorithms are identical with the exception that ls3b checks the misclosure vector during each iteration and discards any observation whose misclosure is greater than some multiple of the Root Mean Square (RMS) value of the misclosure vector.

For N TOA measurements, the misclosure vector is

$$\mathbf{w} = \left(\mathbf{T} - \frac{\mathbf{r}^o}{c} - \mathbf{TOT}^o \right) \qquad (6.15)$$

where $\mathbf{T}$ is a vector of length N containing the TOA measurements,

        $\mathbf{r}^o$ is a vector of length N containing ranges from the N cell sites

to   the approximates coordinates of the telephone,

        $\mathbf{TOT}^o$ is a vector of length N in which each entry is the

                    approximate Time Of Transmission, and

        c is the speed of light.

During each iteration updates to the approximate coordinates and TOT are calculated and used to compute the misclosure vector. Each entry of the misclosure vector, beginning with that of the largest magnitude, is then compared to the misclosure RMS value which is

$$w^{RMS} = \sqrt{\frac{\sum_{i=1}^{N} \mathbf{w}_i{}^2}{N}} \quad . \tag{6.16}$$

If the misclosure for any observation is greater than some multiple (commonly 3) of the RMS value, it is discarded provided that at least three observations remain. For both of the circular trilateration algorithms, the observation covariance matrix is the identity matrix.

Two different LS algorithms were used to implement hyperbolic trilateration. The algorithm lsdifa differences the TOA of the cell site with the highest SNR from that of all other cell sites involved. This is reference differencing as discussed in section 4.2.3.2. and the observation covariance matrix is given by 4.23. Sequential differencing is implemented in the algorithm lsdifsqa. In this case, TOA differences are formed between adjacent entries in the TOA measurement vector. No particular order is observed in the formation of the measurement vector. Sequential differencing results in the observation covariance matrix of (4.28).

In conclusion, the position of each grid point with four or more observations was estimated with one of the four LS algorithms described above. A threshold of four observations was required due to the LS plane intersection method. Only three observations would be required if an initial approximate position is obtained from some other source. If LS converged, the error in the position estimate was recorded.

6.3 Positioning Simulation Results

6.3.1 Uncorrupted Ranges

The stability of the LS algorithms was tested by conducting simulations with uncorrupted ranges. Multipath and TOA estimation noise were generated as described earlier but not added to the LOS ranges. Although not employed, these errors were

generated in order to maintain the same values for the stochastic portion of path loss. This was necessary in order to yield, to the greatest extent possible, the same participating cell sites for each grid point as when range errors are included. In this way a comparison is possible between simulations incorporating multipath and TOA estimation noise and those which do not. Unfortunately, the absence of excess range due to multipath results in shorter ranges for the purposes of path loss calculation. This may result in a higher number of participating cell sites than when ranges are corrupted.

The ls1, lsdifa, and lsdifsqa algorithms were tested with uncorrupted ranges. The ls3b algorithm was not tested in this way since it does not make sense to throw away observations which are perfect. For all three algorithms tested, the estimated position errors and observation residuals were essentially zero (neglecting round-off error). Of the 7211 grid points, 703 were not positioned due to less than four participating cell sites. The maximum number of participating cell sites of the 6508 points which were positioned was 21. The overall mean number of participating cell sites was 8.3 and the standard deviation was 3.9. When the same simulation (i.e. same random number generator seeds) was performed with multipath and TOA estimation noise added to the ranges, 740 points had less than four participating cell sites. The mean number of cell sites was slightly less at 8.1 but the standard deviation was the same. As suspected, for some grid points the lack of excess range resulted in perhaps one or two more participating cell sites. However, as the preceding numbers suggest, the differences are not significant.

6.3.2 Range Corruption by Multipath and TOA Estimation Noise

Simulations were conducted with ranges corrupted by both multipath and TOA estimation noise due to MUSIC. Ten simulation runs were conducted, each with distinct seeds for the random number generators. The various LS algorithms were employed in each of the simulation runs. Each of the simulation runs represents one particular realization of multipath errors, TOA estimation noise, and to a limited extent propagation losses. Therefore, the number and identity of the cell sites seen from any particular grid point will vary somewhat between the simulation runs. Table 6.3 gives pertinent information regarding

each of the simulation runs. Included are the minimum, maximum, mean, and standard deviation of the number of participating cell sites as well as the number of grid points out of the total of 7211 which were not visible from at least 4 cell sites. The statistics regarding the number of participating cell sites do not significantly differ between the simulation runs. This is not unexpected and is indeed desirable. That the simulation runs differ at all is evidenced by the number of grid points with less than four observations. On average, 10% of the grid points had less than four participating cell sites.

Table 6.3   Participating Cell Site Statistics for 10 Simulation Runs

Run #	RNG Seed	Number of participating cell sites				Points with
		Min	Max	Mean	Std	< 4 obs
1	0	0	20	8.1	3.9	740
2	10	0	21	8.2	3.9	718
3	20	0	22	8.1	3.9	724
4	30	0	21	8.2	3.9	761
5	40	0	22	8.1	3.9	749
6	50	0	21	8.2	3.9	724
7	60	0	20	8.2	3.9	702
8	70	0	20	8.1	3.9	752
9	80	0	23	8.2	3.9	695
10	90	0	21	8.1	3.9	771
Average	N/A	0	21.1	8.15	3.9	733.6

The simulation results for hyperbolic trilateration with sequential differencing (lsdifsqa) are given in Table 6.4. Included in this table are the DRMS, horizontal precision, the mean absolute position error as well as the 67th and 95th percentiles for each of the 10 simulation runs of Table 6.3. Also included is the number of grid points which could not be positioned due to divergence of LS. This is expressed both as an absolute number and as a percentage of the total number of grid points which had four or more observations.

Table 6.4   Simulation Results for lsdifsqa , Multipath and TOA Errors

(Hyperbolic Trilateration with Sequential Differencing)

Simulation Run #	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)	Diverging Points
1	529	529	176	118	547	104 / 1.6%
2	790	790	194	119	548	98 / 1.5%
3	442	442	169	119	551	116 / 1.8%
4	676	676	182	120	531	117 / 1.8%
5	626	626	176	120	532	103 / 1.6%
6	829	829	195	118	567	99 / 1.5%
7	841	841	193	121	545	96 / 1.5%
8	496	496	169	114	517	115 / 1.8%
9	710	710	191	120	559	109 / 1.7%
10	879	879	179	118	529	108 / 1.7%
Average	681.8	681.8	182.4	118.7	542.6	106.5 /1.65%

The DRMS and horizontal precision results of Table 6.4 indicate that there is no bias in the positional errors. This is expected since the presence of a bias would indicate some type of error in the process. The variation between the 10 simulation runs is actually quite small. The largest variations occur in the DRMS and horizontal position which varied from 442 m to 879 m. The 67th percentiles vary by only 7 m and the 95th percentiles by 42m. There is also little difference in the number of grid points for which LS diverges. It is evident that the difference between the simulation runs is primarily the number and magnitude of the position estimates which may be described as outliers.

Simulation run 4 was chosen as typical since the results for that run are comparable to the averages calculated over all 10 runs. For run 4, histograms of the multipath errors, TOA estimation errors, the combination of the two, and TDOA residuals are plotted in Figure 6.16. All of the multipath errors are included whereas only those TOA estimation errors and combination of multipath and TOA estimation errors less than 1000 m are included in the corresponding histograms. This corresponds to approximately $\pm 3\sigma$ and 99%

127

of the total errors. In the case of the TDOA residuals, all residuals less than $\pm 5\sigma$ (again 99% of the total) are included in the histogram. The statistics for all of the multipath errors, TOA estimation errors, combination of multipath and TOA estimation errors, and the TDOA residuals are given in Table 6.5. The min and max values correspond to one particular grid point in each case.
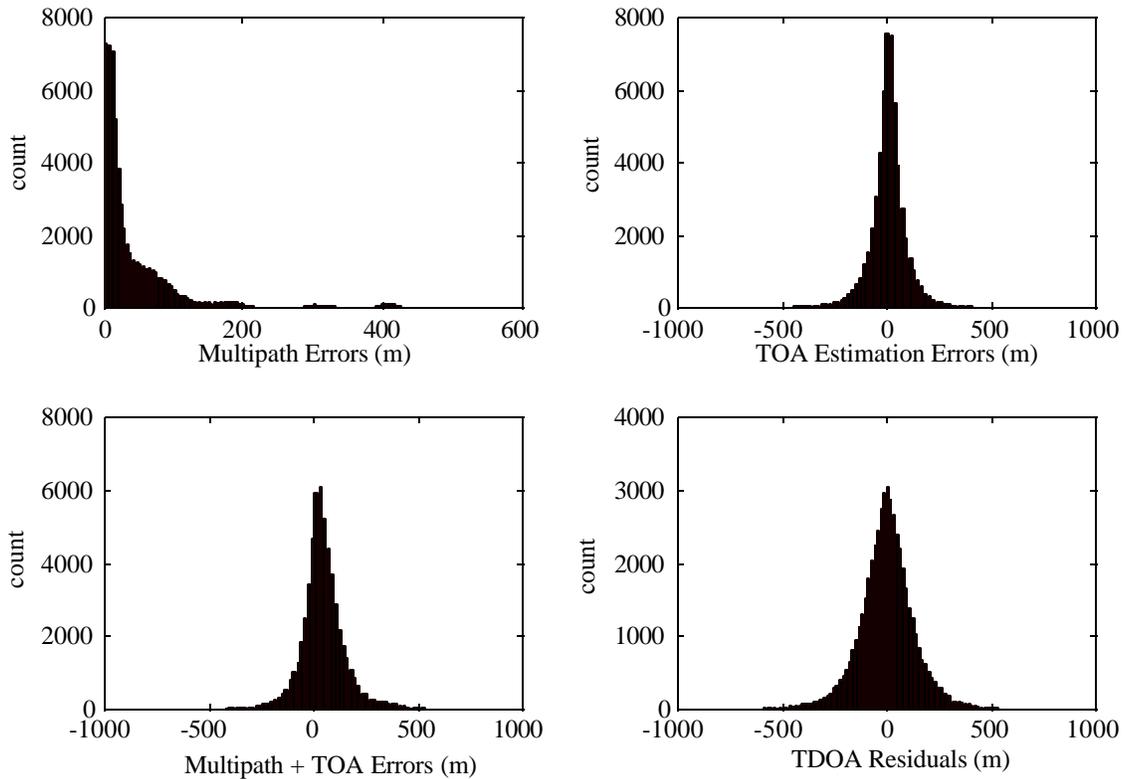
Figure 6.16   Histograms   for   Simulation   Run   4   -   Hyperbolic   Trilateration   with Sequential Differencing

Table 6.5   Statistics for MP, TOA, MP+TOA errors, and TDOA Residuals for   Simulation Run 4

	min (m)	max (m)	mean (m)	std (m)

MP Errors	3e-05	448	42	58
TOA Errors	-31923	44129	6	376
MP + TOA Errors	-31861	44172	47	380
TDOA Residuals	-7563	5864	0.25	152

The values for mean and standard deviation for TOA estimation errors agree with the simulation results reported in Chapter 5. Looking back to Figures 5.7 and 5.9, one sees that a standard deviation of 376 m agrees with those figures when taking into account that the majority of SNRs lie between 13 dB to 20 dB, and multipath phase varies from 0° and 315°. At first glance, the mean TOA error seems small when compared to Figures 5.4 and 5.8. However, the mean error for multipath phases around 180° was negative and therefore pulls the average towards zero. The mean TOA error in Table 6.5 is still positive as expected. The multipath error histogram is also as expected. Combining Figures 6.4 and 6.5 in the same proportion as the number of urban grid points to suburban grid points should yield a result comparable to the TOA estimation histogram of Figure 6.16. Combining the multipath and TOA estimation errors results in slightly biased range errors with standard deviation only slightly greater than that of the TOA estimation errors. The primary effect of the multipath is to bias the range measurements since the standard deviation of the multipath errors are relatively small.

The simulation results for ls1 and lsdifa are very similar to those of Table 6.4 and their individual results, therefore, are not shown. Instead, Table 6.6 compares the averages of the simulation results amongst ls1, lsdifa, and lsdifsqa. As expected the results are very similar. In fact the 67th percentiles are the same. The greatest difference is seen in the DRMS values but the variation is only about 4%. The differences were found to be due to only one or two grid points in each of three of the ten simulation runs. For instance, the only difference between sequential differencing and reference differencing in simulation run 3 was one grid point. LS with reference differencing diverged for this point whereas LS with sequential differencing converged to a very poor solution. The HDOP for this point was 15 and there were four TOA observations. In simulation run number four, there was one grid point which primarily accounted for the difference between sequential differencing and

reference differencing. Although both forms of differencing resulted in convergence to a solution, the error for referencing differencing was orders of magnitude larger than that for sequential differencing. This point again had only four TOA observations and an HDOP of 26.

In general, the three LS algorithms give exactly the same results. In the very few exceptions, the HDOP was very high and the number of observations low. It seems that in these special cases, numerical instabilities cause one algorithm to converge and another to diverge or converge to a wildly different solution. The results herein do not provide sufficient evidence that one algorithm is better than the others in terms of the positioning results. However, between the two hyperbolic differencing methods, sequential differencing is normally favored. Sequential differencing may have an advantage when the geometry is poor. In such cases, the matrix $\left[\mathbf{A}^T \mathbf{C}_1^{-1} \mathbf{A} \right]$ may have a higher condition number for reference differencing than for sequential. Hence, for hyperbolic trilateration, sequential differencing is preferred.

Table 6.6   Average Simulation Results

Positioning Algorithm	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)
Circular (ls1)	694	694	182.7	118.7	542.3
Hyperbolic (lsdifa) Reference Diff.	708	708	182.9	118.7	541.9
Hyperbolic (lsdifsqa) Sequential Diff.	682	682	182.4	118.7	542.6

The principal observation from the results presented thus far is that if the simulations conducted accurately reflect system operation in the field, the FCC specification of 125 m (67%) can be achieved with the cellular infrastructure of Telus Mobility in Calgary and using root MUSIC for TOA estimation. That the DRMS and 67th percentiles do not agree suggest that the error data is not Gaussian but that there are a significant number of large outliers. The largest horizontal positional error encountered in the 10 simulation runs

was 54 km. This was for a point with an HDOP of four. Each of the positioning algorithms was able to converge to a solution over 98% of the time. Therefore, of the total 7211 grid points in the simulations, 12% could not be positioned - 10% due to a lack of observations (less than four) and 2% due to divergence of LS.

The performance of ls3b was also evaluated. As discussed earlier, ls3b employs circular trilateration but discards any observations for which the misclosure exceeds some threshold. Table 6.7 compares simulation results of ls3b with two different thresholds to ls1, circular trilateration in which all observations are retained. The first three simulation runs of Table 6.3 were used to compute the average numbers of Table 6.7. The misclosure thresholds tested were two and three times the misclosure RMS value defined in (6.15).

Table 6.7   Simulation Results for Circular Trilateration with and without Removing Outlier Observations

Positioning Algorithm	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)	Diverging Points
Circular (ls1)	574	574	178.7	118.7	547	106 / 1.6%
Circular (ls3b) Thresh = 3 * RMS	574	574	178.7	118.3	545	106 / 1.7%
Circular (ls3b) Thresh = 2 * RMS	599	599	184.3	118.7	572	152 / 2.3%

Table 6.7 indicates that for these simulations there is no advantage in checking for and discarding outlier observations. The numbers for ls1 and ls3b with a threshold of 3*RMS are virtually identical. When the threshold was lowered to 2*RMS, the DRMS, mean and 95th percentile actually increase although the 67th percentile remains the same. In addition the number of grid points for which LS diverged also increased. It is obvious, therefore, that the multipath and TOA estimation errors introduced in the simulations did not create many observational outliers. When the misclosure threshold was lowered to 2*RMS, it is evident that perfectly valid observations were being discarded. This resulted in poorer

position estimates. It is expected that with actual data, the likelihood of observational outliers is real and that discarding poor observations will improve performance.

6.3.2.1 Effect of Geometry on Positioning Performance

Figure 6.17 demonstrates the effect of geometry on positioning performance for LS, hyperbolic trilateration with sequential differencing. Simulation run 4 was again chosen for this purpose. Figure 6.17 contains the same type of plots as Figure 6.15. Plotted are the DRMS, horizontal precision, and 67th and 95th percentiles for grid points with HDOP less than or equal to the abscissa. Also plotted are the percentage of points, out of the 6333 for which LS converged, and the mean number of participating cell sites for those points with HDOP less than or equal to the abscissa.

Comparing Figure 6.17 to Figure 6.15, one sees a significant improvement in positioning performance when hyperbolic trilateration is used in place of plane intersection. For those points with an HDOP of one or less, the 67th percentile for plane intersection is approximately 1.5 times higher than that of hyperbolic trilateration. The greatest difference, however, is in those errors above the 67th percentile - the outliers. For all points, the 95th percentile for hyperbolic trilateration is one half that of plane intersection. As discussed in section 6.2.7.1, the points responsible for the large DRMS and horizontal precision values for plane intersection are relatively few in number and have poor geometry. This observation suggests that, although the plane intersection and hyperbolic trilateration solutions are both LS, the plane intersection model is more sensitive to poor geometry.

As expected, positioning accuracy improves as the HDOP decreases. There is a large increase in the DRMS, from 100 m to 406 m, when the HDOP increases from 1 to 2. The 67th and 95th percentiles also increase the most in this interval. This is due to a large increase in the number of points with an HDOP less than 2 as compared to points with an HDOP less than 1. The percentage of grid points with an HDOP of 1 or less is 51% and for an HDOP of 2 or less the percentage is 81%. If the percentages are calculated out of the total 7211 grid points as opposed to the 6333 grid points for which a solution was found, the results are 44% and 71%.
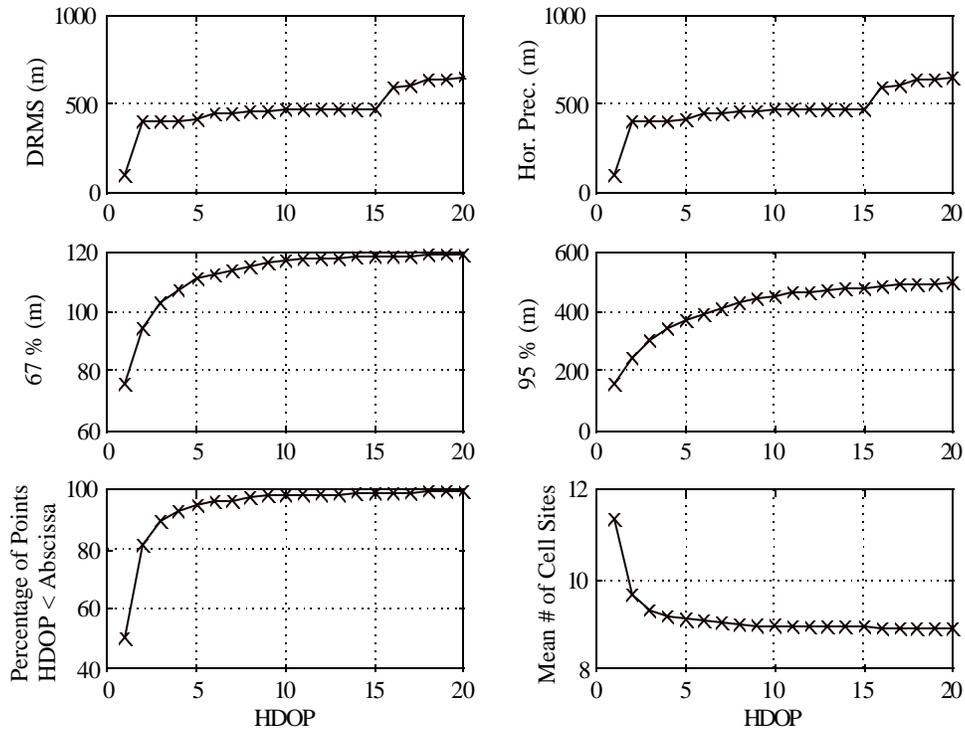
Figure 6.17   Performance Measures, Percentage of Points, and Mean Number
                            of Participating Cell Sites vs. HDOP for LS Hyperbolic
                                      Trilateration   with   Sequential   Differencing
(Simulation Run 4)

Figure 6.17 demonstrates that geometry is a critical factor. To maintain the FCC specification of 125 m (67%) it will be necessary to ensure that the number of cell sites and their positions are such that good geometry will be maintained. This is especially important in suburban and rural areas where the number of cell sites is low due to low channel demand. To maintain good geometry in such areas, it may be necessary to install additional cell sites. These additional cell sites need not be fully functional in terms of offering voice channels to cellular telephones.   Instead, they would exist only for positioning purposes and consequently need very little hardware in addition to the cellular positioning equipment.

6.3.2.2 Effect of Propagation Environment on Positioning Performance

      To determine the effect of propagation environment on positioning performance, the simulation results of Table 6.4 were divided between those points located within the 2 km radius urban zone and those in the 2 - 12 km radius suburban zone. The results are given in Table 6.8.

      Of the 195 points in the urban zone, only 0.2% were visible from fewer than 4 cell sites. This was expected since the city centre contains 6 cell sites in a 2 km radius area. The maximum number of observations for any point in the urban area was 12 and the minimum was 3. Although 4 observations was used as a threshold for the sake of plane intersection, it should be possible to position every point in the urban zone with hyperbolic trilateration since all points were visible from three or more cell sites. In the suburban zone a much larger percentage, 10.4% of the 7016 points, were visible from fewer than 4 cell sites. The minimum number of observations for any particular point was zero and the maximum was 23. Therefore, maintaining a minimum number of observations can be a problem in the suburban zone but not in the urban. The points which were not visible from any cell sites were along the edge of the city.

Table 6.8   Average Simulation Results Divided by Propagation Environment

                      (Hyperbolic Trilateration with Sequential Differencing)

Environment	DRMS (m)	Hor. Prec.(m)	67% (m)	95% (m)	Percentage of Points with < 4 obs	Percentage of Points which Diverged
Urban (< 2 km) 195 points HDOP > 5	1129	1130	327	1895	0.2%	28.5%
Suburban (2-12 km) 7016 points HDOP generally < 3	664	664	116	514	10.4%	0.7%

      Points in the suburban zone fared far better than those in the urban zone in terms of convergence. Although the number of points for which LS diverged is approximately 50 in both zones, the percentage of points which diverge is far greater in the urban zone. This may

be attributed to geometry. Recall from Figure 6.7, that the HDOP in the urban zone was generally 5 or greater whereas for the majority of points in the suburban zone, the HDOP was 3 or less.

Considering the poorer geometry, higher TOA estimation error, and larger multipath effects in the urban zone, it is not surprising that the positioning results are poorer there than in the suburban zone. Considering geometry alone, one would expect the horizontal precision to be approximately 1.7 to 5 times worse in the urban zone than in the suburban zone. In fact, it is 1.7 times worse in the urban zone. The 67th percentile in the urban zone is approximately 3 times greater than that in the suburban. Although the FCC specification is not met in the urban zone, it is still met overall. However, this conclusion is based on interpreting 67% of all cases to correspond to 67% of all locations.

The fact that the horizontal precision in the urban zone is not more than 1.7 times worse than in the suburban zone may be due to the high rate of divergence in the urban zone. Had the quality of the observations or the geometry been slightly better for some of the points for which LS diverged, it is possible that LS would have converged but to a poor solution. Therefore, if solutions would have been obtained for more of the 28.5% of the points which diverged, the DRMS and 95th percentile values would have probably increased and the horizontal precision would have become worse.

6.3.3 Multipath Corrupted Ranges

The ten simulation runs of Table 6.3 were repeated but without TOA estimation noise added to the ranges. The ranges were only corrupted by multipath. However, the multipath errors and participating cell sites are identical to those in the simulations of Table 6.3. Hyperbolic trilateration with sequential differencing was used. The results are given in Table 6.9.

Comparing the results of Table 6.9 with those of Table 6.4, one sees that with multipath only, the 67th and 95th percentiles are less than half of those when both multipath and TOA estimation errors are present. If the multipath models used in these simulations are

realistic, one may conclude that with perfect TOA estimation the FCC specification of 125 m (67%) can be easily met in a city such as Calgary.

Table 6.9   Simulation Results for lsdifsqa, Multipath Errors Only

(Hyperbolic Trilateration with Sequential Differencing)

Simulation Run	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)	Diverging Points
1	562	562	87	52	235	30 / 0.46%
2	273	273	82	54	248	33 / 0.51%
3	382	382	79	51	224	30 / 0.46%
4	1293	1293	92	53	241	26 / 0.40%
5	459	459	82	53	233	32 / 0.50%
6	446	446	85	52	242	25 / 0.39%
7	239	239	76	52	233	23 / 0.35%
8	247	247	76	51	227	34 / 0.53%
9	469	469	85	53	239	30 / 0.46%
10	262	262	80	52	241	29 / 0.45%
Average	463.2	463.2	82.4	52.3	236.3	29.2 /0.45%

Since the ranges in this case are corrupted by multipath only, the results may be compared with those of Morley (1995). Morley reports that for his real network and a dense urban propagation environment, LS resulted in a mean absolute location error of 71 m and a 67th percentile of approximately 72 m. For his moderate urban environment (suburban environment here) the mean and 67th percentile were 11 m and 9 m respectively. This is comparable to the results here - a mean location error of 82.4 m and a 67th percentile of 52.3 m when taking into account that these numbers are for a combination of the two environments.

6.3.4 TOA Estimation Noise Corrupted Ranges

The same 10 simulations were again conducted but this time with ranges corrupted only by TOA estimation noise, not multipath. The results are reported in Table 6.10. In the

simulations in which multipath was included, propagation loss was calculated based on the multipath corrupted ranges. The resulting SNR was the factor which then determined whether a particular cell site participated in the position estimation process or not. By excluding multipath here, the propagation losses are somewhat smaller leading to a greater number of participating cell sites. Therefore, the conditions for these 10 runs are not identical to those of Table 6.3. However, the differences are minor. The maximum, mean, and standard deviation of the number of participating cell sites for these simulations are 22, 8.23, and 3.91 respectively. These values are slightly higher than those of Table 6.3. The greatest differences lies in the number of grid points which had less than four participating cell sites. In Table 6.3 the average is 734 whereas here the average is 682.

Table 6.10   Simulation Results for lsdifsqa, TOA Estimation Errors Only
(Hyperbolic Trilateration with Sequential Differencing)

Simulation Run	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)	Diverging Points
1	402	402	149	99	491	95 / 1.46%
2	545	545	157	98	476	100 / 1.54%
3	440	440	149	101	465	107 / 1.63%
4	1003	1003	161	98	473	95 / 1.46%
5	648	648	152	97	463	120 / 1.83%
6	873	873	177	100	471	96 / 1.47%
7	792	792	157	100	467	86 / 1.31%
8	441	441	147	101	460	94 / 1.44%
9	523	523	148	100	454	90 / 1.38%
10	586	586	159	97	489	81 / 1.24%
Average	625.3	625.3	155.6	99.1	470.9	96.4/ 1.48%

With only TOA estimation error, system accuracy is 99 m (67%) compared to 52 m (67%) with only multipath and 119 m (67%) with both error sources. This is of no surprise considering the histograms of Figure 6.16. The effect of multipath is to positively add to the bias of the TOA estimation errors and only slightly increase the standard deviation. As a result, the addition of multipath to already present TOA estimation errors results in a relatively small increase in the 67th percentile. It may, therefore, be concluded that according to these simulations and the models used, positioning accuracy in terms of the 67th percentile is most effectively pursued by improving the TOA estimation accuracy. The question then is what level of TOA estimation error becomes swamped by multipath?

6.3.5 Various Levels of TOA Estimation Error

To determine how much the TOA estimation error may be decreased until multipath becomes the dominant factor, simulations were conducted with various levels of TOA estimation error. These simulations are identical to those of Table 6.4 with the following exceptions. Whereas the multipath models used are identical to those used in previous simulations, the TOA estimation models are not. In the simulations presented here, one TOA estimation error model is used per simulation run. Different models depending on the propagation environment and a randomly chosen multipath phase are not used. The objective here is to decrease the TOA estimation error and determine when a further decrease in TOA estimation error is of no consequence.

Eight simulation runs, each with a different TOA estimation error model, were conducted. As the previous simulation results indicate, there is little difference in the 67th percentile when the random number generator seeds are varied. Therefore, all eight runs use the same seed. The eight TOA estimation error standard deviation curves used are displayed in Figure 6.18. The mean TOA estimation error for all eight models is zero for all SNRs considered.

The curves were modeled with polynomials and the coefficients used within the simulations to obtain a TOA estimation error standard deviation for any particular SNR. A TOA estimation error was then generated from the Normal distribution with the standard

deviation so obtained and a mean of zero. The TOA estimation error was added to the range already corrupted by multipath. The results of the eight simulation runs are shown in Table 6.11. The TOA Std numbers in Table 6.11 correspond to the labeled curves in Figure 6.18.
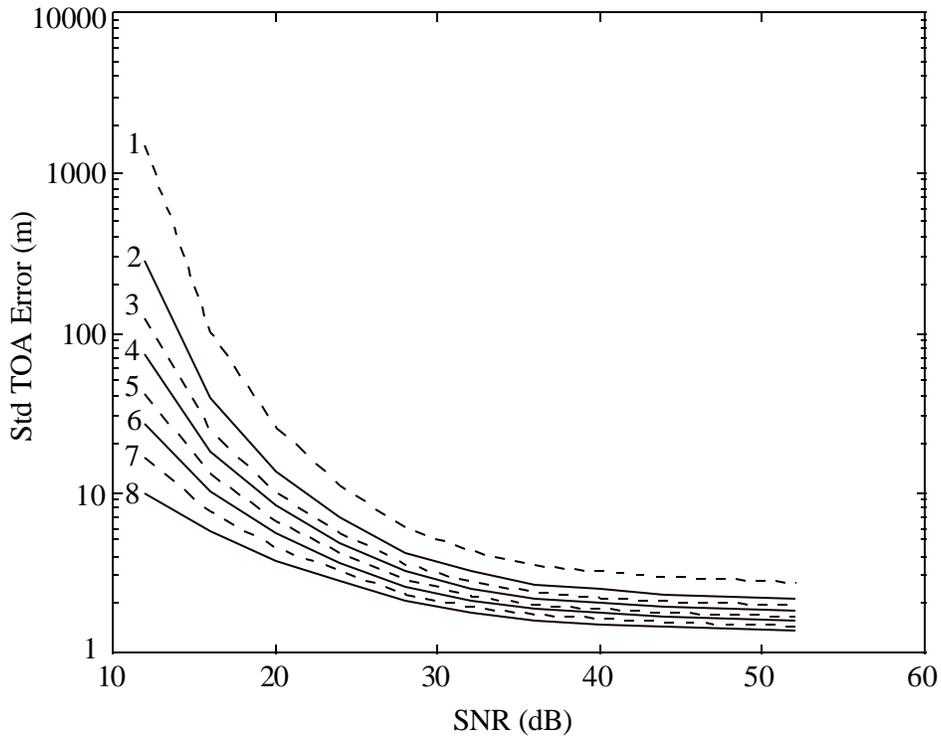
Figure 6.18   Standard Deviation of TOA Estimation Models

Table 6.11   Simulation Results for Multipath and Various Levels of TOA Estimation Error (Hyperbolic Trilateration with Sequential Differencing)

TOA Std Curve	DRMS (m)	Hor. Prec. (m)	Mean (m)	67% (m)	95% (m)	Diverging Points
1	1203	1203	315	208	977	118 / 1.8%
2	502	502	122	81	361	46 / 0.7%
3	315	315	95	63	289	39 / 0.6%

139

4	260	260	87	58	258	34 / 0.52%
5	245	245	82	55	252	31 / 0.48%
6	241	241	81	54	245	31 / 0.48%
7	244	244	80	53	243	30 / 0.46%
8	243	243	79	53	242	31 / 0.48%

As the standard deviation of the TOA estimation error decreases the 67th percentile eventually reaches a floor of about 53 m. This agrees with the simulation results of Table 6.9 which show that with multipath only, the 67th percentile is approximately 52 m. Further decreasing the TOA error standard deviation here, therefore, will not result in a lower 67th percentile. In fact the maximum TOA error standard deviation for curve 8 is already only 10 m.

From Table 6.11, one may conclude that multipath becomes the dominant error source once the TOA error standard deviation reaches that of curve 5 in Figure 6.18, a maximum of 45 to 50 m at 13 dB SNR. From Figures 5.10 and 5.11 one sees that the lowest TOA error standard deviation modeled for 13 dB is approximately 100 m. Therefore, there is still room to improve the TOA estimation process before multipath, as modeled here, becomes dominant. Any reduction of TOA estimation error standard deviation below approximately 50 m will not be significant.

CHAPTER 7

STATIC MULTIPATH FIELD TEST

7.1 Introduction

For the sake of analyzing system performance, it is highly desirable to separate the effect of multipath from that of errors in the Cellocate™ system such as hardware biases and TOA estimation resolution. In the Telus Mobility field tests to follow in Chapter 8, there is a strong probability that multipath varies due to the dynamic nature of the environment. Specifically, the high vehicular traffic in the test area is expected to cause fluctuations in multipath and hence the position results. In an attempt to focus specifically on the Cellocate™ system errors, a test was designed in which multipath and its effects were minimal but more importantly, static. In the complete absence of multipath, all positional error is obviously due to errors in the Cellocate™ system. If multipath is present but static, it will contribute to an overall bias which will be unique to each Cellocate™ receiver. This bias may be modeled and removed from the observations. If this is done properly, the only remaining error source will be the TOA estimation noise due to the Cellocate™ receivers. If the multipath is dynamic, it may be very difficult to distinguish it from the TOA estimation noise which is the very effect we wish to quantify.

7.2 Test Site

The static multipath test was conducted on an acreage near Balzac Alberta, approximately 4 km north of Calgary's city limits on June 7, 1997. A two dimensional plan view of the area is given in Figure 7.1. The black dots indicate the positions of the Cellocate™ receivers. They are labeled according to their position in the field as well as the actual Cellocate™ receiver number. For example, Cellocate™ receiver number 1 was placed at the SW site. The position of the cellular transmitter is indicated by the +. The coordinates of all these points were surveyed with DGPS. The reference station was located on the University of Calgary Engineering building rooftop approximately 18 km away. The occupation time for each position was at least 45 minutes and the data was processed with

Semikin™. The accuracy of the coordinates is better than a few dm which is more than enough for the present purpose.

The shape of the property on which the test was conducted is a long narrow rectangle running east-west. The maximum achievable separation between the receivers on the south side and the receivers on the north was approximately 100 m. The length of the property in the east-west direction was approximately 800 m. However, due to a rise in elevation at an easting of approximately 400 m in Figure 7.1, moving the NE and SE receivers further east would result in a non-LOS path between them and the cellular transmitter. In addition, a house and other buildings are located at the far east end of the property and it was desirable to remain as far away as possible from sources of reflection. Therefore, the distance between the east and west receivers was approximately 400 m. The three-dimensional ranges from the transmitter to each of the four Cellocate™ receivers varied from 187 m to 222 m. The NE and SE receivers were approximately 8 m higher than the NW receiver, SW receiver, and the transmitter.

The geometry of the four receivers and transmitter is quantified by the DOP numbers in Figure 7.1. An HDOP of 1.95 for the transmitter location is the best possible given the geometry of the receivers. The EDOP and NDOP reflect the fact that the receiver east-west separation is approximately four times the north-south separation. One therefore expects that the standard deviation in the northing of the estimated transmitter position will be four times that of the easting.
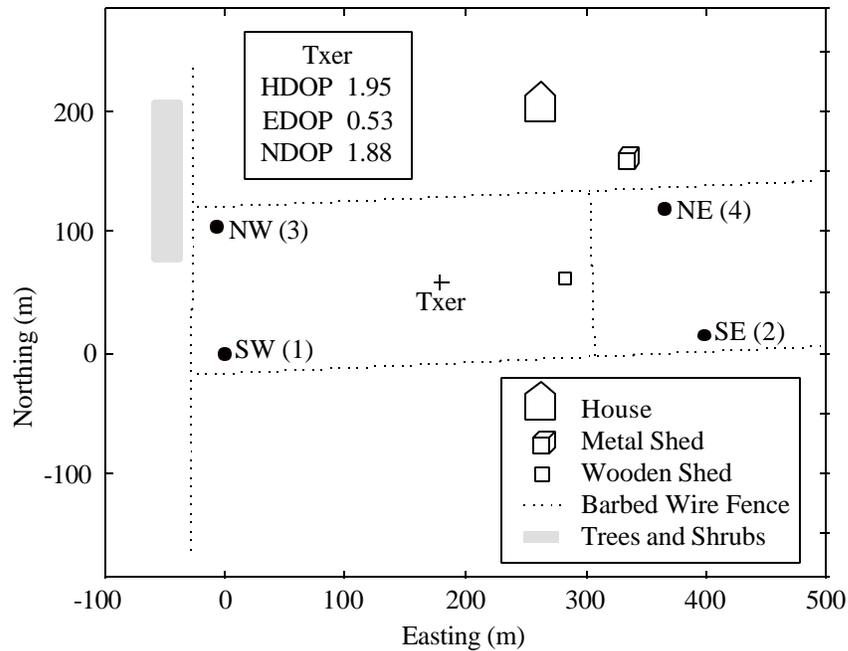
Figure 7.1   Test Site for Static Multipath Test

There were four objects in the immediate vicinity of the test site which may be sources of signal reflection and diffraction. Due to the short distances involved and the similarity in elevation of both the transmitter and receivers, refraction should be negligible. The major sources of reflection are the metal shed and house on the neighboring property to the north. One would expect that the NE receiver in particular would be affected by these two sources of multipath. The wooden shed near the transmitter did not block the LOS path from the transmitter to the NE and SE receivers. It may, however, cause a certain amount of diffraction which in turn also causes multipath. Foliage, such as the trees in the NW corner of the field, may cause diffraction, reflection, and absorption. Since the trees do not lie in the path between the NW receiver and the transmitter, diffraction and absorption are not considered threats. However, they may cause some reflection of the signal. In any case, since all of these objects were stationary, any multipath created by them should be static. During the test there was little movement in the test area. Data was collected over a span of

approximately one hour and during that time the author walked between receivers for a total of about 5 minutes. The only other human movement was a pickup truck which drove the length of the field from east to west in the adjacent field to the north. This occurred very near the end of the one hour and only lasted about 2 minutes.

7.3 Equipment Setup

The Cellocate™ and GPS receivers were set up at each of the four receiver stations in the configuration discussed previously. Omni-directional RF antennas were used for the Cellocate™ receivers. The GPS receiver at each station was fixed to the surveyed coordinates for that station. Both the GPS receivers and Cellocate™ receivers were allowed to run for at least one hour prior to the collection of data. This was more than sufficient for the receiver clock model in the NovAtel GPS receivers to stabilize and for the Cellocate™ receivers to reach a stable operating temperature.

To facilitate testing without the interference of other cellular telephones in the area, a reserved reverse control channel was used. The frequency of this channel was 835.02 MHz. Since cellular telephones lock onto whichever control channel is strongest, a transmitter was built in order to simulate the transmission of the registration message on this frequency. Had an actual cellular telephone been used, it would have been necessary to tune the Cellocate™ receivers to the strongest control channel in the area. The Cellocate™ receivers would then have received registration messages from any active cellular telephone in the vicinity. This would require sorting through all the received messages and using the mobile identification number which appears after the precursor in the registration message to identify the correct message. Although necessary in an operational system, this is less than convenient during testing. Using a transmitter allowed transmission on the reserved control channel frequency. The Cellocate™ receivers could then be tuned to this frequency and only received from the transmitter since no cellular telephones would use this channel.

The transmitter simulated the same message and modulation scheme as an analog cellular telephone. It was programmed to transmit the registration precursor every 5 seconds in order to collect as much data as possible in a minimum amount of time. The transmission

power level was approximately 0.5 W which is typical for modern hand held telephones. Both the transmitter antenna and the Cellocate™ antennas, with the exception of the SW antenna, were at ground level. The antenna at the SW station needed to be raised about 1 m off the ground in order for the Cellocate™ receiver to consistently receive the transmitter's message. This was due to the sensitivity of that particular Cellocate™ receiver, not its location. The received SNR at each of the receivers was unknown. Although the limit of the Cellocate™ receiver at the SW station was approached, it is expected that the other receivers still had sensitivity to spare. During developmental work it was found that the Cellocate™ receivers with the omni-directional antennas had a range of approximately 2 km in a relatively built-up area. In the operational field tests to follow in Chapter 8, the low noise amplifiers and directional antennas of the Telus Mobility cell sites are expected to increase the range of the Cellocate™ receivers.

7.4 Test Results

7.4.1 Data

The data collected consists of 502 epochs common to all four Cellocate™ receivers and spanning 56 minutes. The data is not evenly spaced due to some missing epochs. Since the transmission rate was once every 5 seconds, there should be 670 epochs of data in 56 minutes. Occasionally one or more of the receivers did not receive a transmission at all or the received record was incomplete. In other cases the received data was corrupted to the extent that the GPS time stamp from the coarse correlation process was not close enough to the time stamps of the other receivers for that particular transmission. The GPS time stamp for each received record of a particular receiver was compared to all other record time stamps of the other receivers. Those time stamps within 1 ms of each other were assumed to correspond to the same transmission. On occasion all four receivers would receive the same transmission but one or more of the receiver's data was corrupted such that the time stamp was outside the 1 ms threshold. In that case, a position fix was not attempted since that particular TOA observation would contribute a large amount of error. If more redundancy

were available such an observation could be discarded as described in section 6.2.7.2 and a position fix for the epoch still obtained.

7.4.2 Cellocate™ Biases

Each Cellocate™ receiver attempts to estimate the TOA of the cellular signal impinging on its antenna. The TOA estimates must be synchronized in order to estimate position. The GPS receiver at each site provides this synchronization. When a Cellocate™ receiver locks onto the Barker word the GPS receiver is pulsed and records a time stamp which is later referenced by the TOA estimation software. Before the GPS receiver is pulsed however, there are numerous delays within the RF equipment and the Cellocate™ receiver. These delays are due to the antenna cable, RF hardware within the Cellocate™ receiver as well as delays in the DSP software. In the operational tests of Chapter 8, there are additional delays due to the front end equipment in the Telus Mobility cell sites. For each of the receivers and cell sites these delays will differ. The combined effect of these delays will henceforth be called the site bias.

Because the site biases are on the order of the propagation times involved, it is necessary to accurately determine them in order to estimate position from TOA. Without accurate knowledge of the biases, the LS position estimation algorithm diverges. Measuring the absolute site biases is difficult. A simplified range equation for any particular site may be written as

$$p_i = \left(TOA_i + Rx_i - TOT + T_B + \boldsymbol{e}_i \right) \cdot c \qquad (7.1)$$

where p_i is the geometrical distance between the transmit antenna and the antenna of site i;

TOA_i is the estimated time of arrival at site i;

Rx_i is the time bias at site i;

TOT is the actual time of transmission at the transmitter;

T_B is the time bias in the recorded *TOT*,

$\boldsymbol{e}_i$ are unmodeled errors such as multipath and resolution of the

TOA estimation process, and

c is the speed of propagation.

Although the *TOT* at the transmitter may be obtained by pulsing a GPS receiver, it is impossible to pulse the GPS receiver at the exact time the signal is transmitted. Hence there is a bias, T_B in the *TOT* estimate. If this bias were known and propagation errors in $\boldsymbol{e}$ could be minimized, it would be possible to reasonably estimate the site biases by transmitting from a known position. Since T_B is unknown, it is impossible to accurately measure the absolute site bias Rx. However, the bias differences may be determined. Since the *TOT* and T_B are common to all sites receiving the signal, they cancel when TOA differences are formed. If the known range differences are subtracted from the TOA differences, the result is the bias differences and the combination of the unmodeled errors.

Since the *TOT* cancels when TOA differences are formed, it is not necessary to measure it. For the test under consideration here, it is assumed that the contribution of ε_i in (7.1) is negligible; differencing TOAs and subtracting the known range differences is therefore believed to result in accurate bias differences. In the Telus Mobility tests to follow this cannot be assumed and the bias differences are measured in another manner.
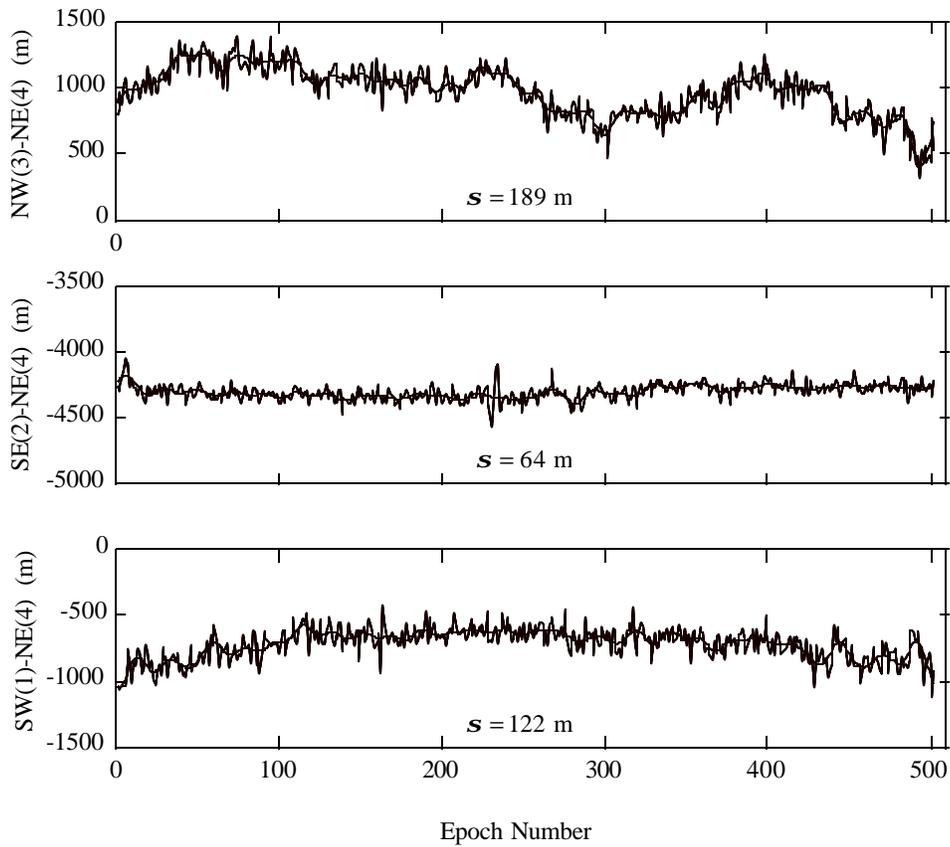
Figure 7.2   Raw and Lowpass Filtered Bias Differences (NE Reference)

After the common epochs between all four Cellocate™ receiver's were found, each receiver's data was processed as described in section 5.4. The result is a TOA at each of the four sites for every common epoch. Corresponding TOAs of any two sites were then differenced and the corresponding known difference in range to the transmitter subtracted. The results are given in Figures 7.2 through 7.5. In each figure a different site was chosen as the reference which is subtracted from the other three. Consequently, there are three plots in each figure. Note that the vertical axis scaling in all four figures is identical. In each plot there are two data sets shown. The noisy data set is the raw TOA differences with the range

differences subtracted. The second is a lowpass filtered version of the first. The standard deviation of the raw data is also given in each plot. The cell sites are again identified by their field position and Cellocate™ receiver number.
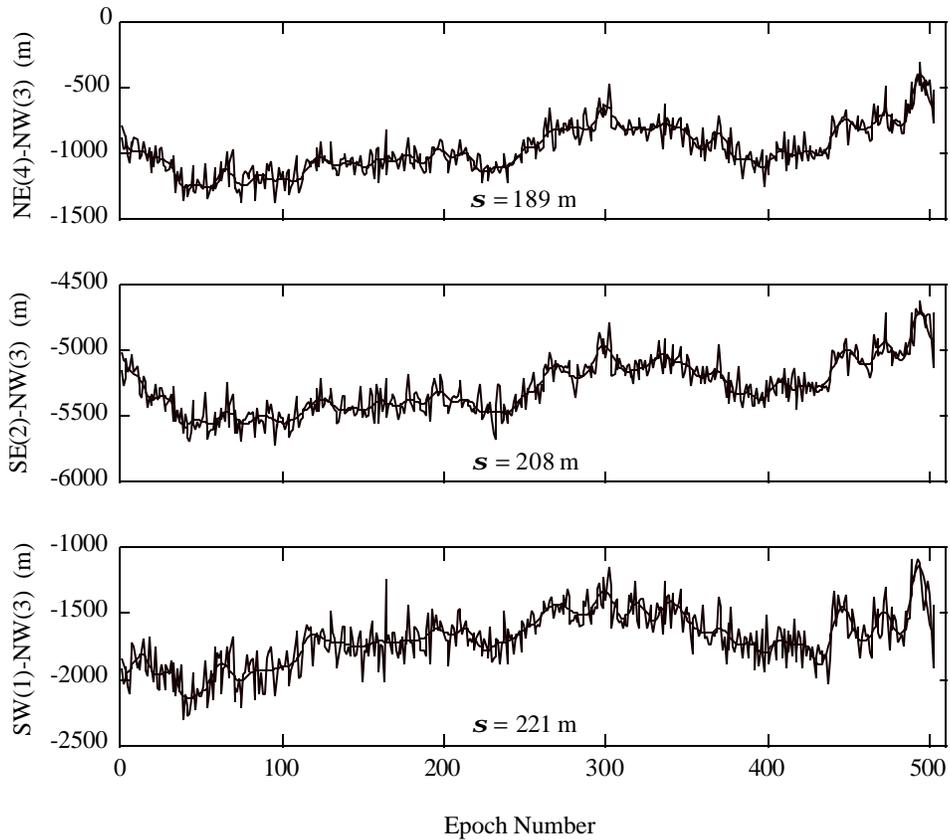
Figure 7.3   Raw and Lowpass Filtered Bias Differences (NW Reference)

There are two distinct characteristics in the raw data. It is obvious that there is short-term white noise. This noise may be considered the random error in the Cellocate™ system (the TOA estimation noise). There is also a long term drift in the bias differences. This is illustrated by the variation in the filtered bias differences. The level of white noise appears to remain relatively constant throughout the one hour of data. To quantify this

noise, the long term drift was removed by subtracting the lowpass filtered curve from the raw data. The standard deviation of the resulting data is a measure of the level of white noise in the system. These standard deviation values are given in Table 7.1 along with the standard deviations of the raw data and the lowpass filtered data. The standard deviation of the noise varies from 49 m to 113 m. Four of the six values are within 4 m of 79.5 m, the average noise standard deviation. If the long term drift in the bias differences is eliminated, and given an HDOP of 1.95 for the transmitter location under consideration, a horizontal precision error of approximately $1.95 * 79.5 / \sqrt{2} = 110$ m may be expected.
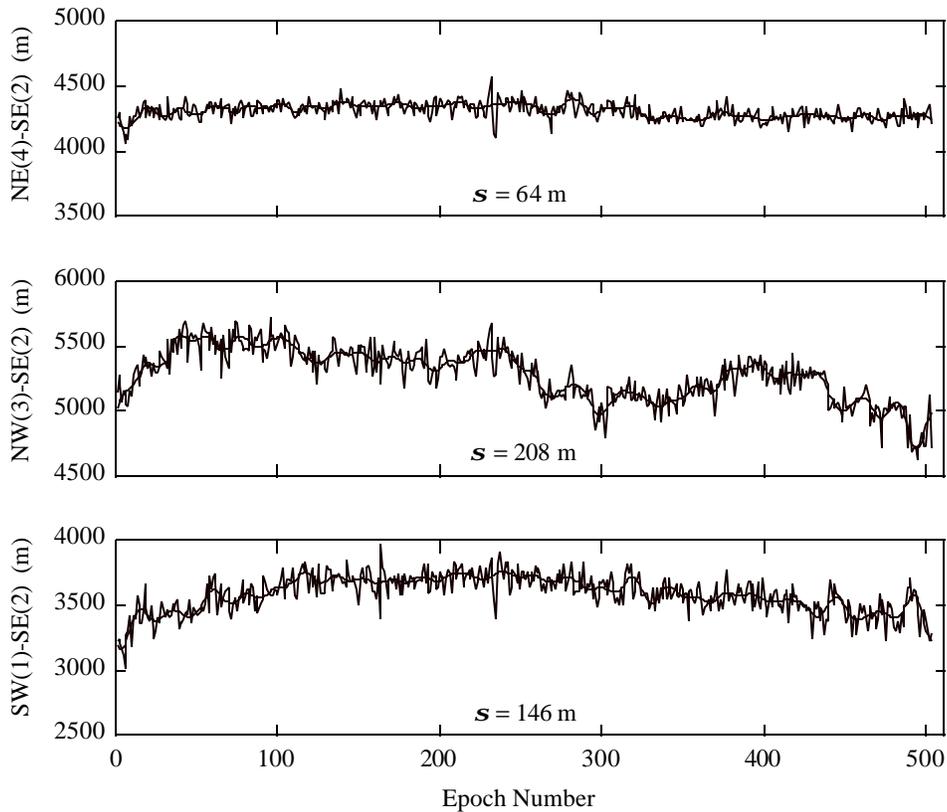
Figure 7.4   Raw and Lowpass Filtered Bias Differences (SE Reference)

The short term white noise does seem to vary from site to site. The standard deviation of the bias difference between the NW and SW sites is significantly larger than the mean whereas the standard deviation of the bias difference between the NE and SE sites is significantly smaller. With the data collected it cannot be determined with certainty whether this is due to the Cellocate™ receivers themselves or the sites at which they were located. This could only be determined by repeating the test with the receivers sequenced through the different sites.
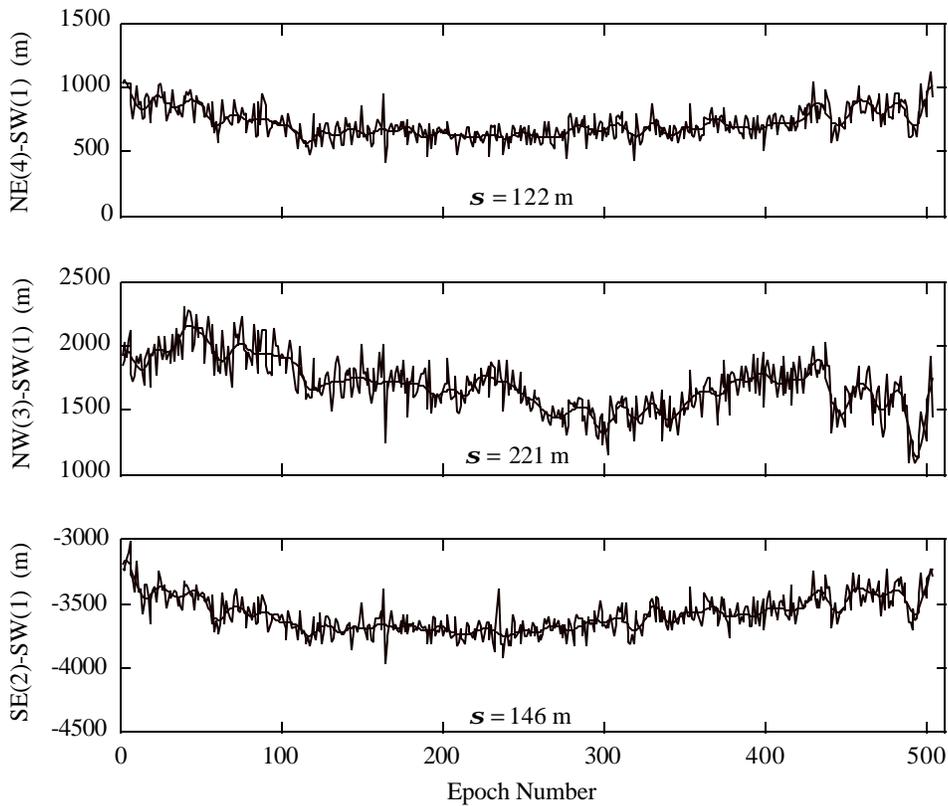
Figure 7.5   Raw and Lowpass Filtered Bias Differences (SW Reference)

It is suspected, however, that the white noise level is a function of both site location and receiver sensitivity. The NE and SE sites were 8 m and 6 m, respectively, higher than the transmitter. The NW and SW sites were both about 2 m lower than the transmitter. In addition the NE site is 2 m higher than the SE site and the standard deviation of any bias involving the NE site is 5 to 6 m less than that of the corresponding bias difference involving the SE site. One may infer that received SNR is a function of the receiver elevation with respect to transmitter elevation. From Table 7.1 one may also conclude that receiver 3 (NW) displays a higher noise level than the other three receivers. Similar results from the operational field tests will confirm this.

The long term bias drift is also site or receiver dependent. There is a distinctive long term drift pattern in each bias difference involving the NW site. The same may be said for any bias difference involving the SW site. However, the NW pattern dominates the SW in the NW-SW bias difference. By far, the NE-SE bias difference is the best behaved. According to Table 7.1, the standard deviation of the long term drift of any bias involving the NW site is at least four times greater than that of the NE-SE bias difference. Bias differences involving the SW site have a long term drift standard deviation at least two times greater than that of the NE-SE difference.

Table 7.1   Standard Deviations of Bias Differences

	S_{raw} (m)	$S_{filtered}$ (m)	$S_{raw-filtered}$ (m)
NW(3) - NE(4)	189	173	76
NW(3) - SW(1)	221	190	113
NW(3) - SE(2)	208	191	82
NE(4) - SW(1)	122	95	76
NE(4) - SE(2)	64	41	49
SW(1) - SE(2)	146	121	81

As opposed to the bias difference white noise, the long term drift is more than likely due to temperature. The group delay of the RF filters currently used in the Cellocate™

receivers may drift in the order of microseconds as temperature changes. Some of the bias differences in Figures 7.2 through 7.5 drift up to 1000 m over the one hour. This corresponds to approximately 3 μs. Since the Cellocate™ receivers were in close proximity of one another and all more or less equally exposed to the elements, it is unlikely that they experienced radically different temperature gradients. In fact one would expect larger temperature drifts for the eastern receivers since they were located atop a small rise and yet they exhibited the least drift. One concludes that the RF filters in the western receivers were more sensitive to temperature. This may also contribute to differences in the white noise levels.

New versions of the RF circuitry in the Cellocate™ receivers include custom designed filters with group delay less dependent upon temperature. It is anticipated that with improved filters, the biases will be relatively constant and the need for periodic calibration will be eliminated. In that case, the biases could be measured over a sufficient length of time and averaged to obtain a single value. This approach was used for the data here but the results, as presented in the next section, are poor. On the other hand, for the operational tests conducted with the Telus Mobility sites, the measured biases were more constant and average values yielded reasonable position results even for data collected on different days.

7.4.3 Plane Intersection Position Results

The Plane Intersection method was investigated for the purpose of determining its suitability for providing an initial position. For any position estimation algorithm, the Cellocate™ biases must be removed from the TOA observations. The bias differences analyzed in the previous section were therefore subtracted from the appropriate TOA differences and the results used in the plane intersection algorithm. Since plane intersection is non-iterative, a solution is obtained for each epoch of data.

In the first case considered, the average of each bias difference was calculated from the full hour of data. The position estimates of plane intersection using these mean bias differences are shown in Figure 7.6. The figure is similar to Figure 7.1. Shown are the positions of the four Cellocate™ sites and a + indicating the true position of the transmitter.

Due to the large scale of the map, the Cellocate™ sites and transmitter position all lay on top of one another. The small dots indicate the instantaneous transmitter position estimates from the same data from which the average bias differences were derived. Also given in the figure are the values for the DRMS, horizontal precision, 67th percentile, and 95th percentile. Notice that the northing and easting are given in kilometres. At this scale it is impossible to distinguish between the four sites as well as the transmitter site. Obviously, the position estimates are grossly in error. It is certain that given the distances between the sites and transmitter are in the order of hundreds of metres, LS will not converge with an initial position estimate in error by thousands of kilometres.

DRMS   = 3614 km
Hor. Prec. = 1916 km
67%   = 3189 km
95%   = 4736 km
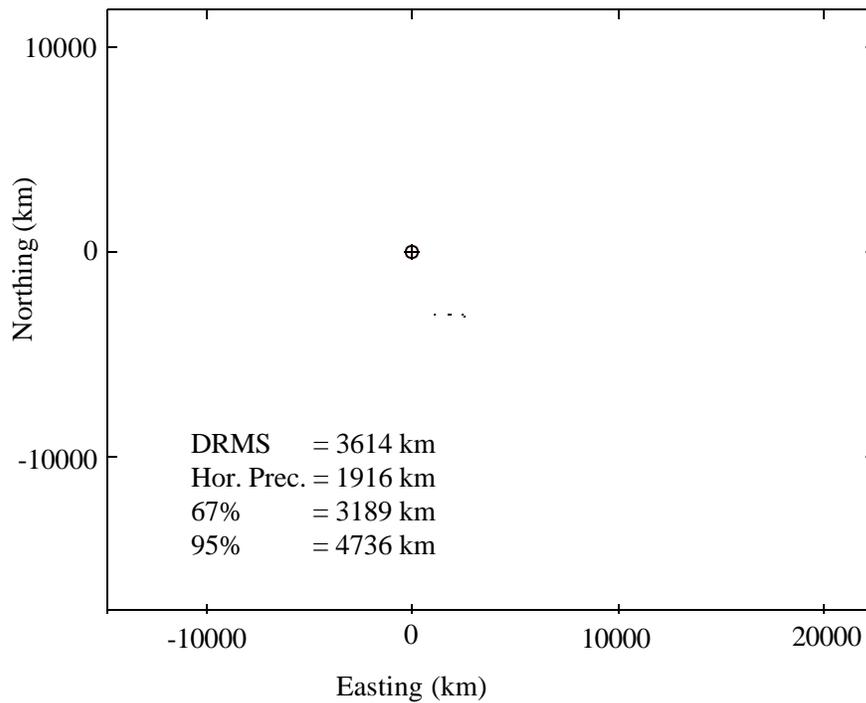
Figure 7.6   Static   Multipath   Position   Results   Using   Plane Intersection   and Average Site Bias Differences

Because the plane intersection method worked well in the simulations of Chapter 6, it is suspected that the site biases are to blame for the poor performance here. To test this hypothesis, the lowpass filtered bias differences of Figures 7.2 through 7.5 were used instead of average bias differences. Whereas the same average bias differences are used for each epoch of data, the lowpass filtered bias differences are used on an epoch by epoch basis. That is, at each epoch, the lowpass filtered bias difference for that particular epoch is subtracted from the TOA difference at that epoch. If instead of the lowpass filtered bias differences the raw bias differences are used, the position estimate should be perfect. Using the lowpass filtered bias differences is akin to assuming that the long term bias drifts may be eliminated in an operational system.

The position estimation results for the case of lowpass filtered site bias differences are given in Figure 7.7. As expected, there is a significant improvement in the positioning performance. Two position estimates were excluded from the plot in order to maintain a small scale. The two estimates excluded were 2.5 km and 15.5 km in error. These two estimates are, however, included in the performance measures. According to the 67th percentile, plane intersection should yield a position estimate accurate enough to be used as an initial position for LS. This will be confirmed in section 7.4.4.4.

DRMS = 715 m
Hor. Prec. = 712 m
67% = 56 m
95% = 216 m
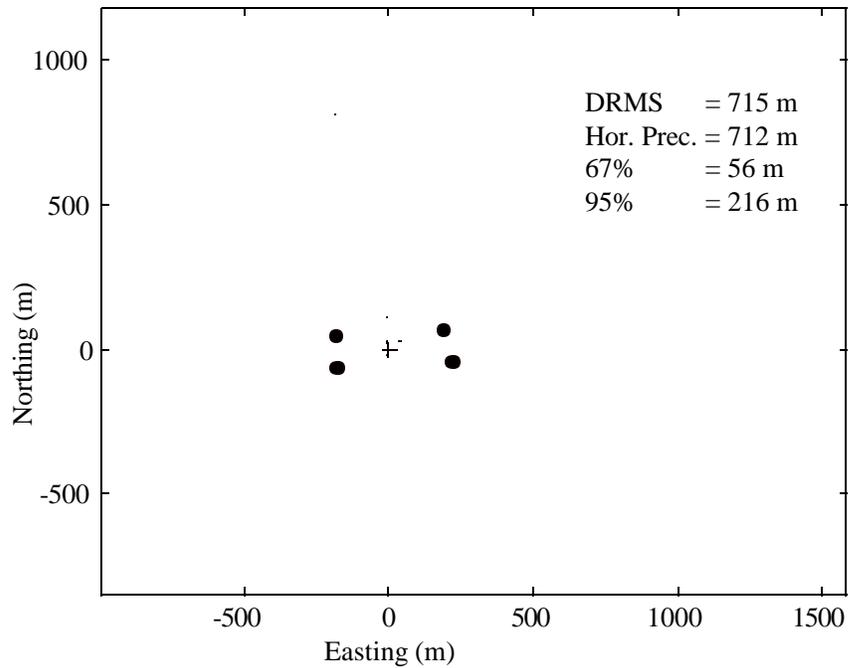
Figure 7.7   Static   Multipath   Position   Results   Using   Plane Intersection   and   Lowpass   Filtered   Site   Bias Differences

7.4.4 Least Squares Position Results

The data collected was also processed by LS to produce position estimates for the transmitter. Because only bias differences are available, either hyperbolic trilateration or circular trilateration with bias differences must be employed. Since the two algorithms give identical results but circular trilateration produces residuals for the individual TOA observations, the results presented are from circular trilateration with bias differences.

The circular trilateration algorithm used chooses one of the four sites as reference and differences that site's TOA with the other three. In terms of position results the choice of reference is insignificant; the DRMS, horizontal precision, 67th percentile, and 95th percentile for the case of the NE site as reference are within 2 m of those of the case when

the NW site is chosen as reference. However, when average biases are used, there is a difference in the estimated values of the transmitter bias; the standard deviation of the transmitter bias is 74 m in the case of the NE reference and 101 m for the case of the NW reference. Since the transmitter bias parameter is the range from the transmitter to the reference site, it is reasonable that the effect of the ill-behaved bias of the NW site (as compared to that of the NE site) will appear in this parameter when the NW site is the reference. If the lowpass filtered biases are used, the choice of reference station becomes even more insignificant. In any case, for the results to follow, the NE site was used as the reference since according to the bias results it seems to be the best behaved.

7.4.4.1 Least Squares Position Results with Average Biases

Figure 7.8 shows position estimates for the case when average bias differences are subtracted from the TOA differences. The initial position used to begin LS for each data epoch was 100 m to the north and 100 m to the east of the true transmitter position. Although the accuracy in terms of DRMS and 67th percentile is good, these statistics are based on only 272 of the 502 epochs of data. For the other 230 epochs, LS diverged. Again this is due to the use of average site bias differences which contain long term drift. It will be shown later, that using lowpass filtered bias differences results in a much higher rate of convergence.

The effect of the drift in the bias differences appears in the estimated parameters. Figure 7.9 shows the unknown parameter estimates as a function of time. Plotted are the errors in easting position, northing position, and the transmitter bias. As discussed in Chapter 4, this transmitter bias should equal the range from the transmitter to the site chosen as the reference. In this case, that range is 197 m. Also contained in each plot are the mean, standard deviation and RMS for the error in each parameter estimate. Comparing to Figure 7.2, one sees a high degree of correlation between the easting position error and the bias difference involving the NW site. The northing and transmitter bias also exhibit, in a more limited sense, the same pattern.
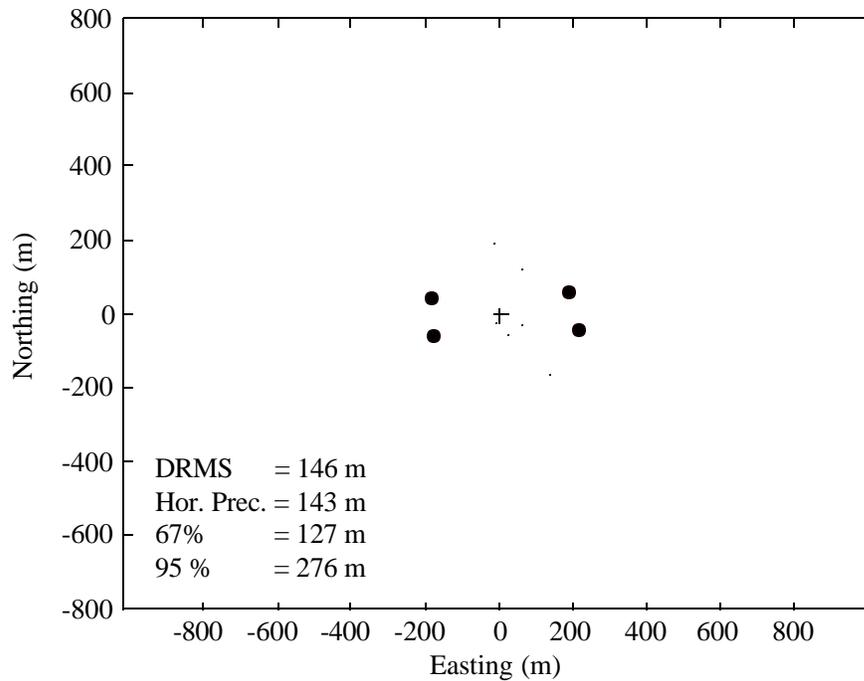
Figure 7.8   Static Multipath Position Results Using Least Squares
and Average Site Bias Differences

The observation residuals are also affected by the bias drift. Consider Figure 7.10. Plotted are the observation residuals as a function of time. The mean and RMS are given for each. Recall that although the observations consist of each site's TOA (including NE) minus the NE site's TOA, the residuals obtained are the same as those that would be obtained by circular trilateration with TOA observations. The same familiar pattern and degree of correlation can be seen in all of the residuals.
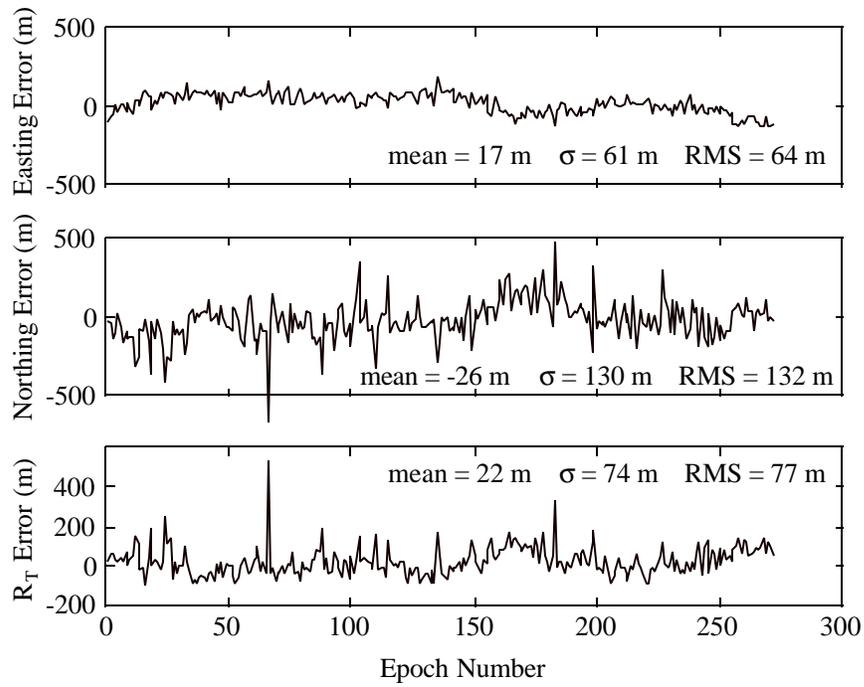
Figure 7.9   Static Multipath Parameter Estimates Using Least

Squares and Average Site Bias Differences

159
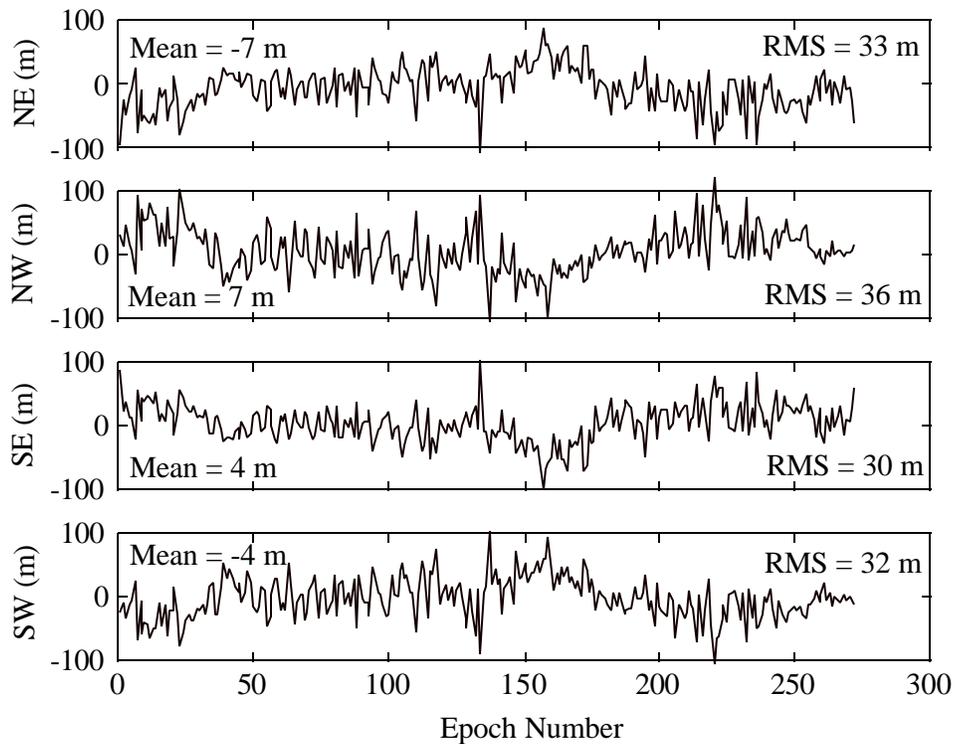
Figure 7.10   Static Multipath Residuals from Least Squares
and Average Site Bias Differences

In order to fully appreciate the effect of the long term drift in the bias differences, the data was processed again but without the NW site data. This leaves only three observations and therefore no redundancy. Although there are no residuals, it is still instructive to view the position estimates and convergence rate. Figure 7.11 contains the position results. Note the increased DOP values due to the loss of one site. Of the 502 epochs of data, LS now converges for 420. Other than the 67th percentile, all of the accuracy measures have increased from the case when all four sites are used. This is due to two reasons. Obviously the loss of redundancy and increase in DOP degrades the position solution. Secondly, the removal of the effect of the NW bias causes LS to converge for 150

epochs for which it did not converge with the inclusion of the NW data. Undoubtedly, the data at some of these 150 epochs is of marginal quality. Inclusion of the position estimates at these epochs results in a larger DRMS, horizontal precision and 95th percentile.
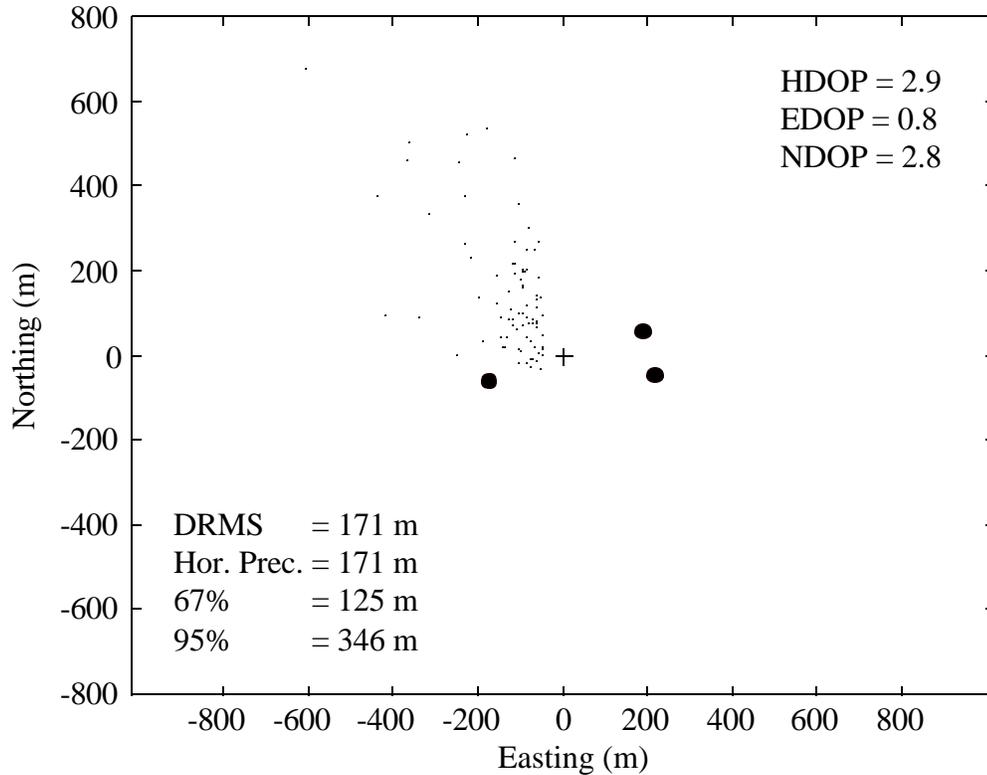
Figure 7.11   Static Multipath Position Results Using Least Squares and Average Site Bias Differences (without NW Site)

The parameter estimate errors as a function of time are shown in Figure 7.12. Again the mean, standard deviation, and RMS of the estimation error is given in each plot. Compared to Figure 7.9, one sees that the parameter estimates no longer resemble the NW bias difference. With the removal of the NW data from the solution, the next most influential bias difference becomes apparent, the SW-NE bias difference (see Figure 7.2).

From these results one obviously concludes that drift in the Cellocate™ site biases directly affects the position estimate. It is highly desirable to eliminate any drift in the biases. As discussed earlier this drift is most likely temperature dependent. Stable operating temperatures and higher quality RF filters should mitigate this problem.
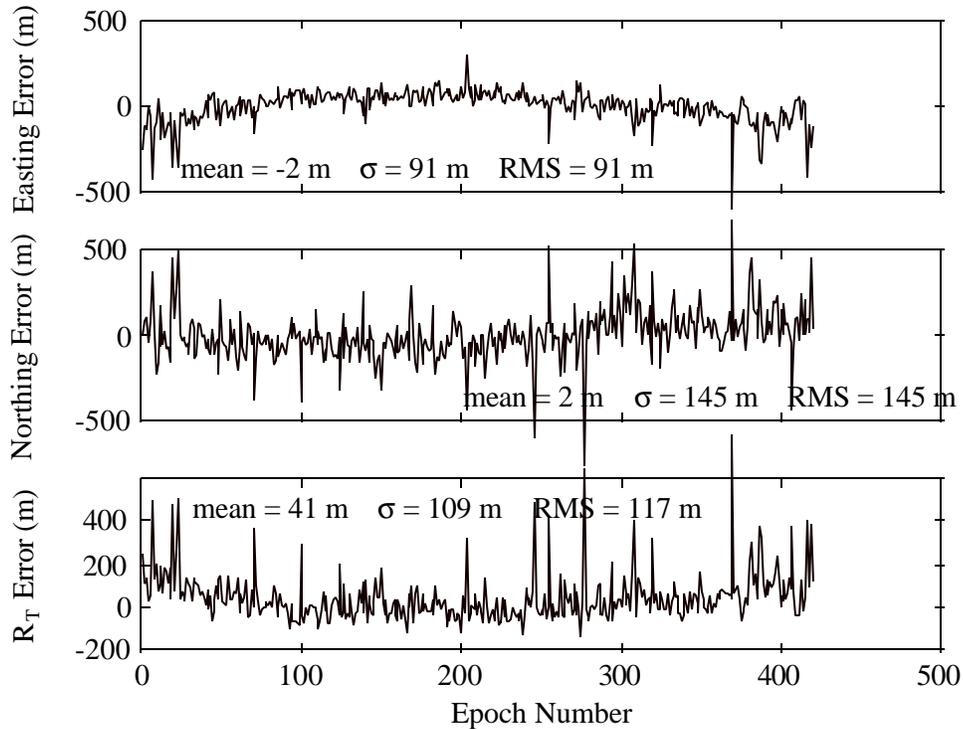
Figure 7.12   Static Multipath Parameter Estimates Using Least Squares   and Average Site Bias Differences (without NW site)

Note that the residuals of Figure 7.10 are highly correlated with one another. In particular, the NE and SE residuals are negatively correlated as are the NW and SW residuals. This is due to the high degree of symmetry in the geometry and only one degree

of freedom. For every epoch the covariance matrix of the observation residuals was calculated. A typical unscaled result is

$$\mathbf{C}_{\hat{\mathbf{r}}} = \begin{bmatrix} 0.274 & -0.308 & -0.210 & 0.244 \\ -0.308 & 0.346 & 0.236 & -0.274 \\ -0.210 & 0.236 & 0.161 & -0.187 \\ 0.244 & -0.274 & -0.187 & 0.217 \end{bmatrix}. \qquad (7.2)$$

Note that the off-diagonal terms are as large as the diagonal terms. From $\mathbf{C}_{\hat{\mathbf{r}}}$ it is obvious that the NE residual is highly negatively correlated with the NW and SE residuals and positively correlated with the SW residual. Correlation coefficients were calculated between the four sets of residuals and are presented in Table 7.2.

Table 7.2   Residual Correlation Coefficients (4 Sites only)

	NE	NW	SE	SW
NE		-0.933	-0.986	0.940
NW			0.888	-0.991
SE				-0.917

Since the $\mathbf{C}_l$ matrix used was the identity matrix, the redundancy numbers of the observations are equal to the diagonal terms of (7.2). The average redundancy number of the 272 convergent solutions is 0.25 which is significantly less than 0.5 which is considered the lowest typical redundancy number of a well designed geometry (Mackenzie, 1985). Therefore, each of the observations is poorly controlled and as a result, very little individual observational error will show up in the corresponding estimated residuals.

7.4.4.1.1 Simulations

In a working system, it is highly desirable that poor observations be detected and removed if possible. If the observations are fully controlled, the residuals may be used as an indicator of observation quality. In the case under consideration here, the observations are

poorly controlled. This suggests that large errors in any one observation will not be detectable in that observation's residual but spread amongst all residuals. To demonstrate this, data was simulated for the geometry of the static multipath test. Gaussian noise was added to perfect TOAs for each of the four sites given the transmitter position. The standard deviation of the noise on the NE site was set at 100 m. The noise standard deviation at the remaining three sites was set to 1 m. Five hundred epochs of data was simulated; of those LS converged for 432. The observation residuals are shown in Figure 7.13.
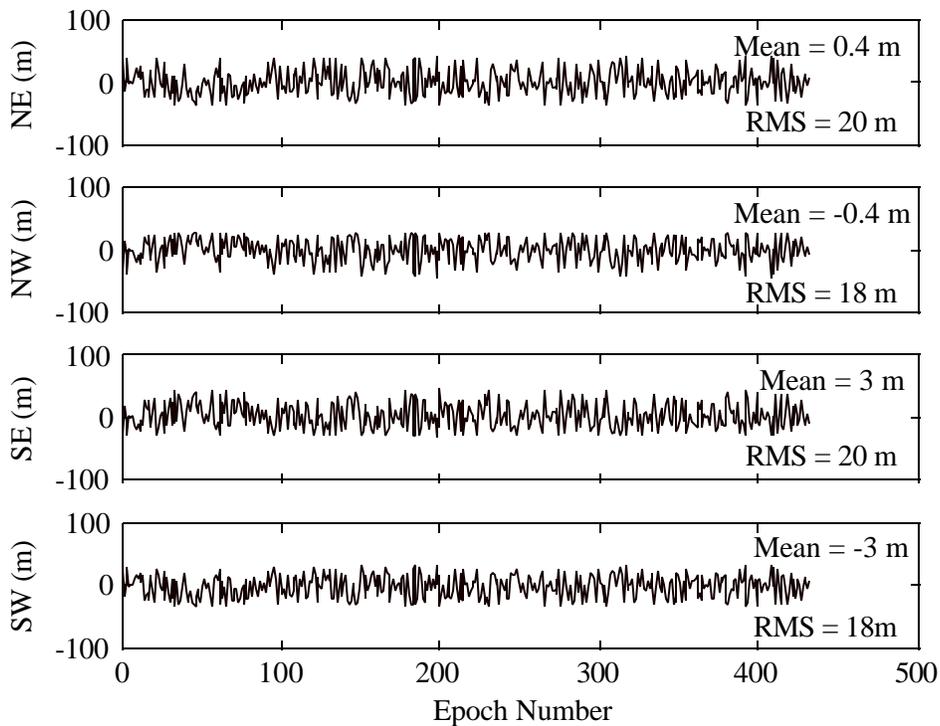
Figure 7.13   Static Multipath Residuals for Four Sites and

Simulated Data (σ_{NE} = 100m, σ = 1m for all others)

Again the residuals are highly correlated; correlation coefficient magnitudes range from 0.98 to 0.997. The residual RMS for each site is either 18 m or 20 m. It is, therefore,

not readily apparent from the residuals that the NE observations are 100 times noisier than the other three. A typical redundancy matrix is

$$\mathbf{C}_{\hat{r}}\mathbf{C}_{l}^{-1} = \begin{bmatrix} 0.299 & -0.267 & -0.278 & 0.247 \\ -0.267 & 0.238 & 0.249 & -0.220 \\ -0.278 & 0.249 & 0.260 & -0.230 \\ 0.247 & -0.220 & -0.230 & 0.204 \end{bmatrix}. \qquad (7.3)$$

and the average observation redundancy number is of course 0.25 as before.

Four additional sites were added to increase the redundancy. The geometry of the network now appears as in Figure 7.14. The four original sites are NE, NW, SE, and SW. The four additional sites are labeled e1, e2, e3, and e4. The transmitter position is again indicated by the + and the position results by the small dots. Note the improvement in the DOP values due to the addition of receiver sites.
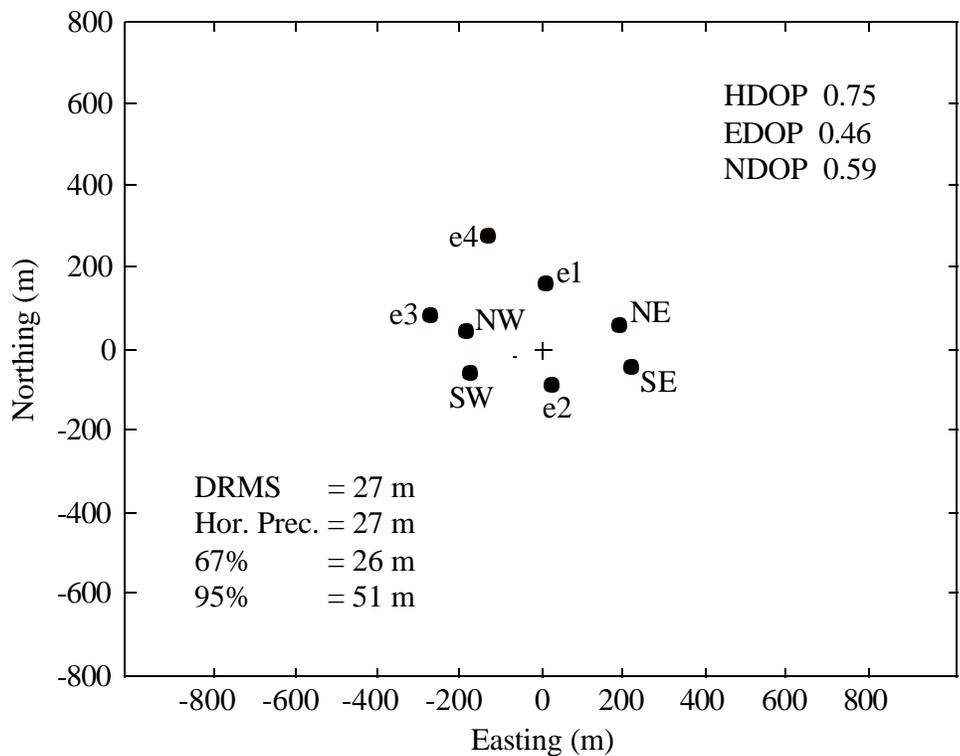
Figure 7.14   Static Multipath Position Results for Eight Sites and

Simulated Data ($\sigma_{NE} = 100$ m, $\sigma = 1$m for all others)

For 500 epochs of simulated data, LS converged for all epochs. The average observation redundancy number was 0.625, a significant improvement over 0.25. The observation residuals for the position results of Figure 7.14 are shown in Figure 7.15a and Figure 7.15b. Correlation coefficient magnitudes range from 0.836 to 0.997. Obviously some residuals remain highly correlated due to geometry and the location of the position fixes. The lowest correlations are found between site e2 and sites NW, SE, SW, and e3.
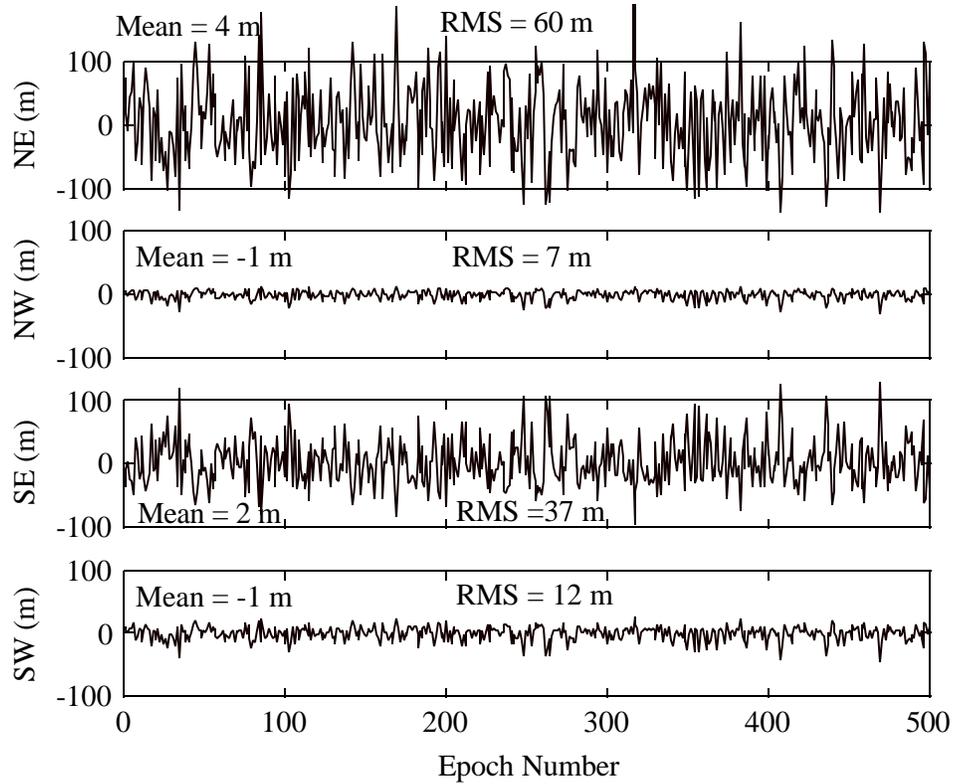
Figure 7.15a   Static Multipath Residuals for Sites NE, NW, SE, SW with
Simulated Data for 8 Sites ($\sigma_{NE} = 100$m, $\sigma = 1$m for all others)

Figure 7.15 clearly illustrates how, with an increase in redundancy number, the observation residuals more accurately reflect observational error. From Figure 7.15a and 7.15b one could reasonably infer that the NE observations contain more noise than the others. However, one may also be led to believe that the SE and e1 sites are also noisier. This example only helps to demonstrate the need for redundancy if the residuals are to be used as indicators of individual observation quality.

167
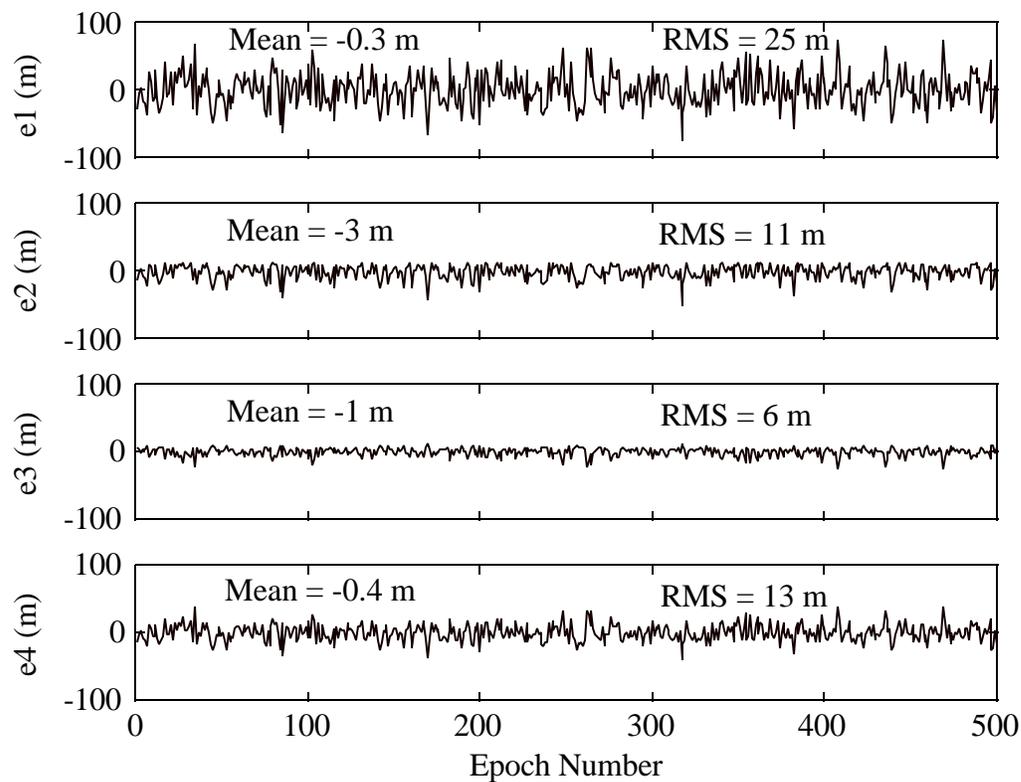
Figure 7.15b   Static Multipath Residuals for Sites e1, e2, e3, e4 with Simulated

Data for 8 Sites ($\sigma_{NE} = 100$ m, $\sigma = 1$m for all others)

7.4.4.2 Internal Reliability and Statistical Testing

As discussed in Chapter 4, it is possible to detect observation blunders by statistically testing the residuals. This was done with the simulated data just investigated. The first step was to ensure that the observation covariance matrix was properly scaled. Because the level of noise added to the perfect observations was known, the covariance matrix could be properly scaled and the a priori variance factor, σ_o^2, set to 1. The a priori

variance factor $\sigma_o^{\,2}$ was still tested with the χ^2 test. The significance level, α, chosen for this example was 0.05 which results in a two-tailed confidence interval of 95%.

Five hundred epochs of data were again simulated and Gaussian noise was added to the observations of all four original sites. The standard deviation of the noise added to each exact TOA observation was 1m. LS converged for all 500 epochs. The χ^2 test on the a priori variance factor passed for all but 27 epochs and the average a posteriori variance factor was 1.01.

The marginally detectable error was explained in Chapter 4 and is repeated here for convenience. For a given α and β, and therefore δ_o, redundancy number r_i, and observation standard deviation σ_i, the MDE of the i$^{\text{th}}$ observation is,

$$\nabla_o l_i = \sigma_i \, \frac{\delta_o}{\sqrt{r_i}} \,. \tag{7.4}$$

For a significance level, α, of 0.05 and a β (probability of Type II error) of 0.10, the noncentrality parameter, δ_o, is 3.24. Given a standard deviation of 1 m and average redundancy number of 0.25, the MDE for the simulated observations is 6.48 m. A blunder of 6.5 m was added to each TOA for the NE site, for epochs 200 to 299. The residuals at every epoch were then statistically tested with (4.42) in order to detect blunders. For "in-context" testing (N=4), the test statistic,

$$n(0,1)_{1-\frac{\alpha}{2N}} \tag{7.5}$$

is equal to 2.5.

Figure 7.16 is a plot of the standardized residuals $\hat{r}_i^{\,*}$. The horizontal line corresponds to the test statistic threshold of 2.5. The standardized residuals for the NE site generally exceed the test statistic of 2.5 from epochs 200 through 299 and thereby detect a blunder for those epochs. Unfortunately, the standardized residuals for the other three sites

are virtually identical to those of the NE site. The observation containing the blunder cannot be identified, again due to low redundancy.
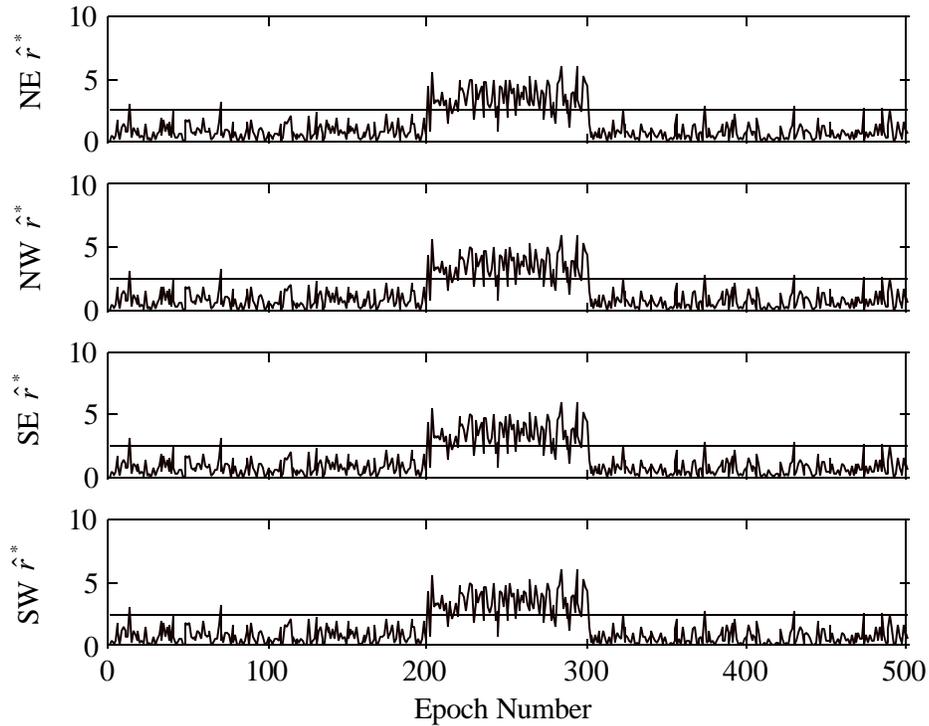
Figure 7.16   Standardized Residuals based on 4 Sites
(Blunder of 6.5 m on NE Site)

Data was then simulated for sites e1, e2, e3, and e4 in addition to the four original sites. The geometry is, therefore, as in Figure 7.14. The observation noise standard deviation was again 1 m for all sites and a blunder of 6.5 m was added to the NE observations for epochs 200 to 299. Standardized residuals were calculated for all eight sites and the results plotted in Figures 7.17a and 7.17b. Note that for eight observations the test statistic threshold is 2.74. It is now obvious that the blunder is in the NE observation. The SE and e1 standardized residuals also indicate a blunder but do not consistently exceed the threshold to the extent that the NE residuals do.
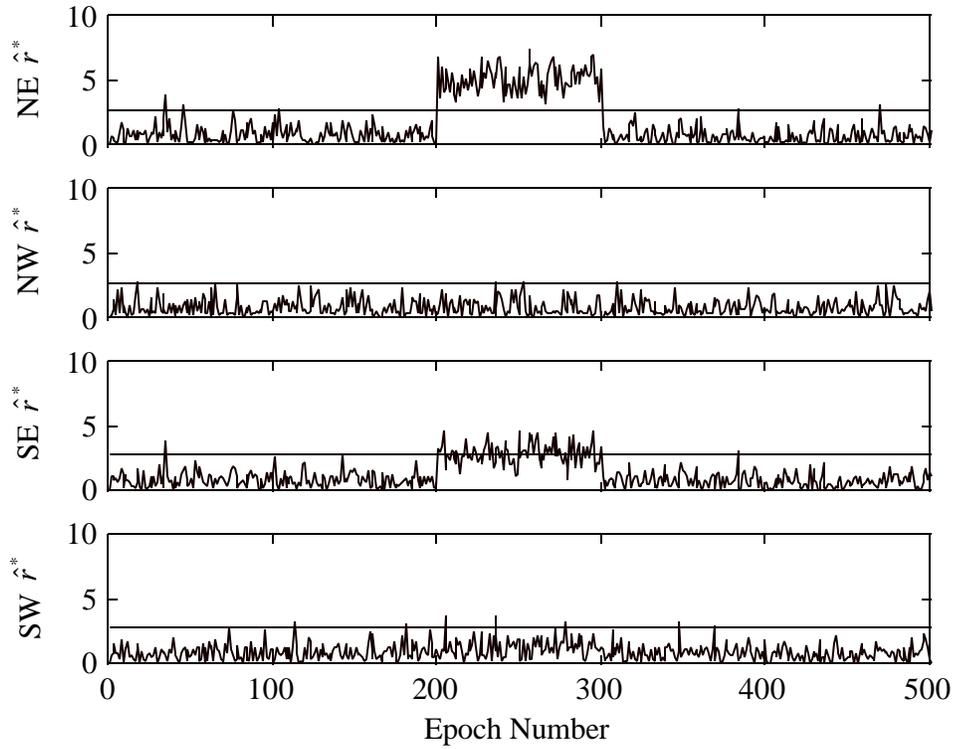
Figure 7.17a   Standardized Residuals for Sites NE, NW, SE, SW

based on 8 Sites (Blunder of 6.5 m on NE Site)
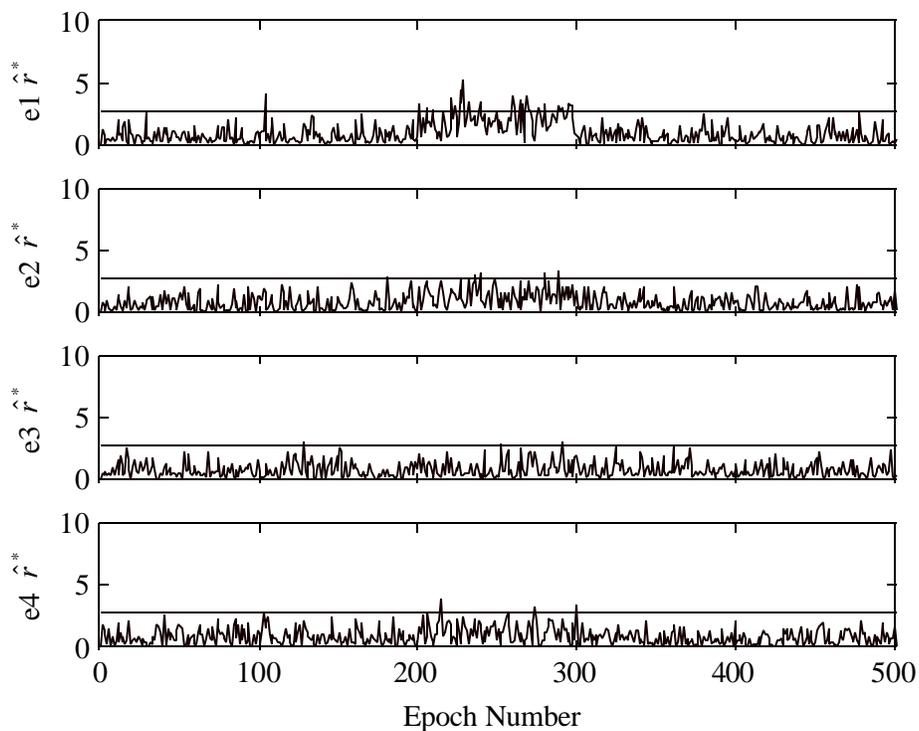
Figure 7.17b   Standardized Residuals for Sites e1, e2, e3, e4

based on 8 Sites (Blunder of 6.5 m on NE Site)

7.4.4.3 Least Squares Position Results with Lowpass Filtered Biases

In section 7.4.4.1 average biases were used. In this section the lowpass filtered bias differences of Figures 7.2 are used. Use of the filtered bias differences removes the long term drift from the observations. For this test, the remaining error source should be the white noise of the TOA estimation process. Any static multipath will cause a bias which will be included in the bias differences. This bias is removed, however, when the filtered bias differences are used. This leaves TOA estimation noise as the major source of error. As explained earlier, the main purpose of this static multipath test was to estimate the positional error, due to Cellocate™ system noise, in a multipath free environment.

With the use of filtered bias differences, the position results are as given in Figure 7.18. The DOP values are the same as in Figure 7.1 since the same four sites are used. In this case LS converged for 425 of the 502 epochs. The 67th percentile is now 105 m, the best so far. The horizontal precision of 118 m is close to the 110 m predicted in section 7.1.4.2. As expected, the position results are essentially unbiased since the DRMS and horizontal precision agree. Upon comparing the DRMS and 67th percentile, one may conclude that the position errors are also more or less Gaussian.
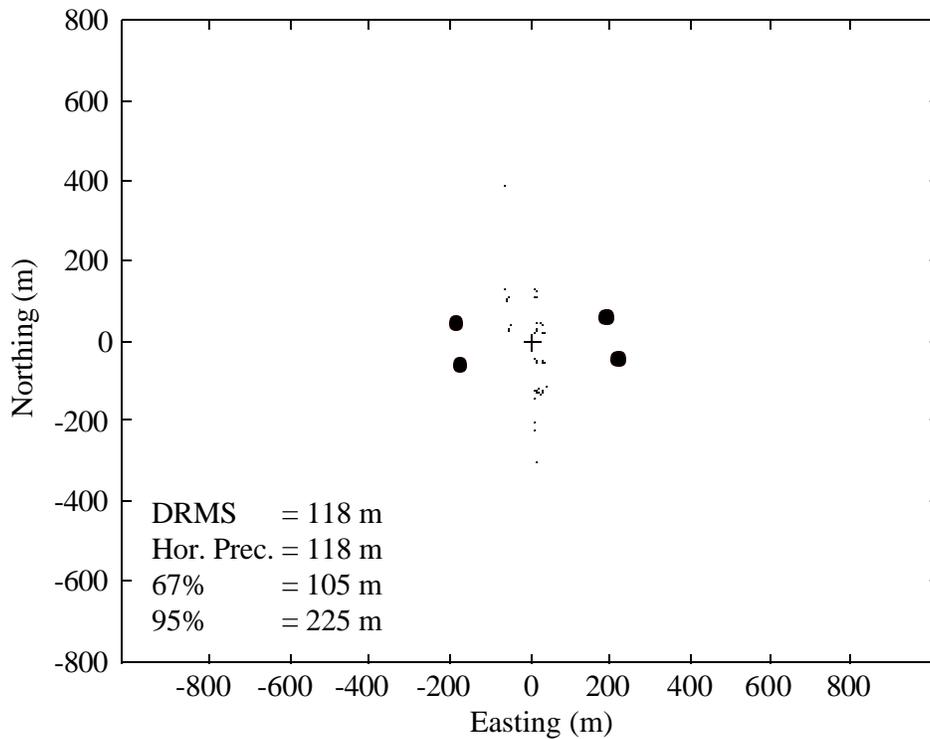
Figure 7.18   Static Multipath Position Results Using Least Squares
and Lowpass Filtered Site Bias Differences

The parameter estimation errors are presented in Figure 7.19. Note that compared to Figures 7.9 and 7.12, there is no detectable long-term drift in the parameter estimates here. The standard deviation of the easting is approximately one quarter that of the northing as expected upon comparing the EDOP and NDOP. Both the easting and northing are essentially unbiased since their means are about one tenth of their standard deviations. It is not clear why there is a large bias in the transmitter bias parameter. It is definitely not caused by a bias in the filtered bias differences. The mean differences between the raw site bias differences and the filtered versions are zero. One is tempted to place the blame on multipath since the bias is positive. However, static multipath is accounted for by the filtered bias differences.
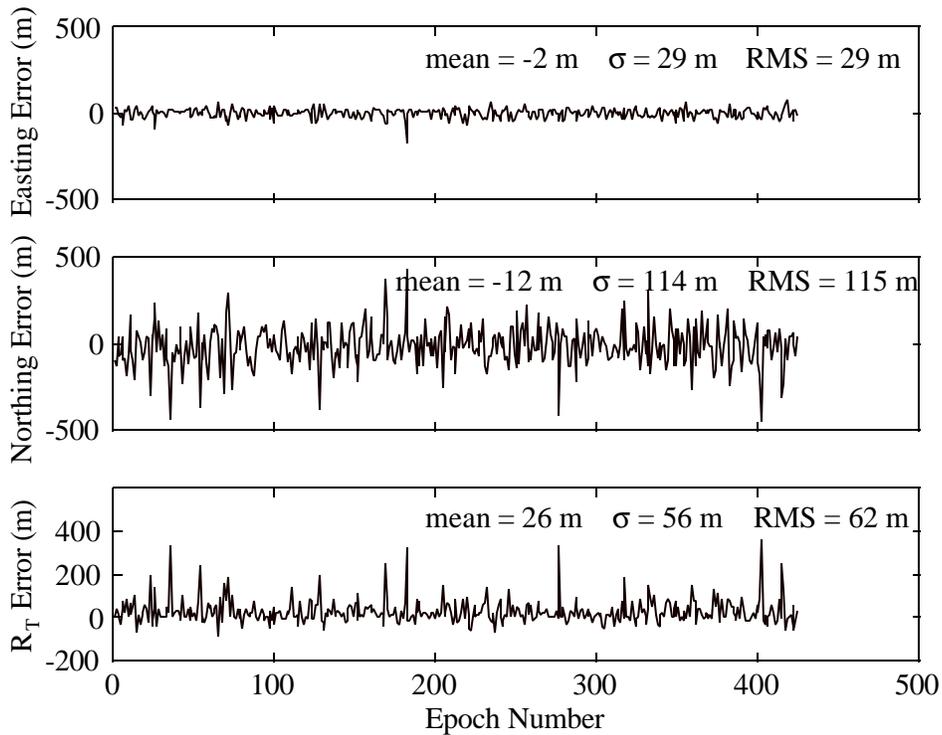
Figure 7.19   Static Multipath Parameter Estimates Using Least Squares
and Lowpass Filtered Site Bias Differences

Figure 7.20 shows the observation residuals. All traces of the long-term drift of the site bias have disappeared. The NW residuals have a significant bias which may be related to the bias in the transmitter bias parameter.

The objective of this field test was to estimate the position error due to the Cellocate™ system in the absence of multipath and compare this to the simulation results of Chapter 6. The results of Table 6.10 predict that in the absence of multipath, TOA estimation error in the Cellocate™ system will produce a horizontal position error of 99 m (67%) for the Telus mobility cellular network in Calgary. To compare this result to the 105 m (67%) result from this field test, one must ensure that the geometry and received SNR are comparable. According to Figure 6.14, just over 80% of the test points in the simulations had an HDOP of two or less and the rate of increase in the 67th percentile slows considerably as the HDOP increases above two. Therefore, taking into consideration that the field test is only one realization whereas the simulations are an ensemble average of many realizations, the results presented in this chapter tend to confirm the simulations of Chapter 6. With good geometry (HDOP ≤ 2) and over short LOS distances, the Cellocate™ system is currently capable of a positioning accuracy of approximately 100 m (67%). Unfortunately the SNR corresponding to short distances is unknown at this time.
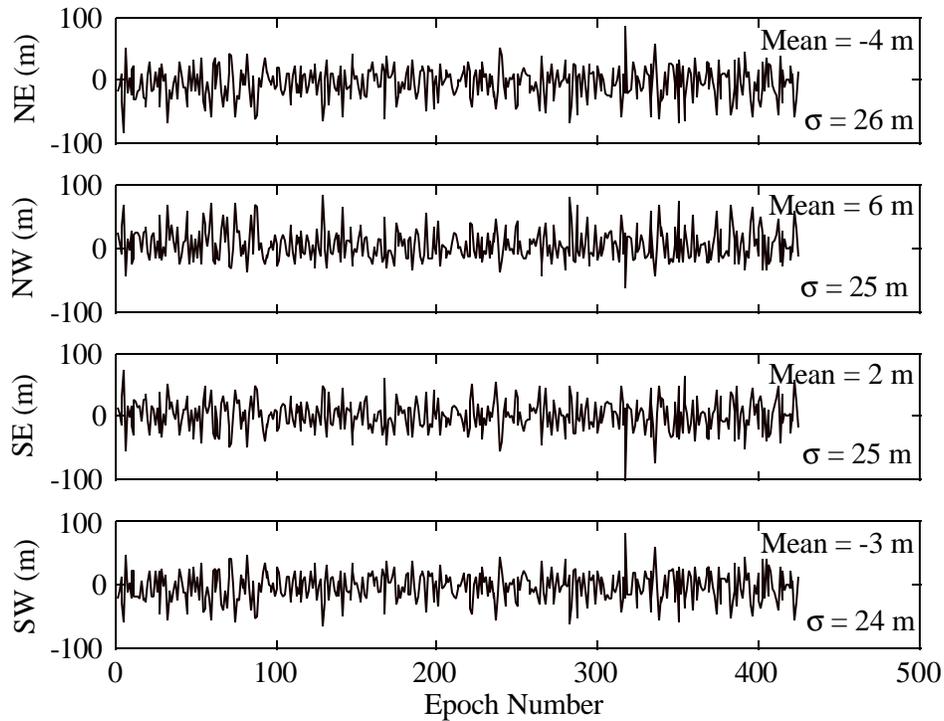
Figure 7.20   Static Multipath Residuals From Least Squares
and Lowpass Filtered Site Bias Differences

7.4.4.4 Effect of Averaging

The results presented so far are all instantaneous; that is no averaging was performed on the position estimates. In the presence of zero-mean Gaussian noise only, averaging the position results over time will improve the position standard deviation by averaging out the noise. The improvement in the standard deviation of the average positional error may be predicted with the equation

$$\sigma(\overline{X}) = \sqrt{\frac{N-n}{N-1}} \, \frac{\sigma}{\sqrt{n}} \qquad (7.6)$$

where $\sigma(\overline{X})$ = standard deviation of the average positional error,

   N = number of position fixes from which to sample,

   n = number of position samples which are averaged,

   σ = standard deviation of the total number of position fixes, N.

The term $\sqrt{(N-n)/(N-1)}$ is a correction factor used whenever the population sampled is finite (Neter et al., 1978). The data set under consideration here contains 425 elements and since sample sizes up to 90 are used, the correction factor is employed.

Figure 7.21 shows the improvement in horizontal improvement due to averaging for the position results of Figure 7.18. Each time interval of the abscissa represents a sample size. For each sample size moving position averages were calculated across the entire 425 position fixes. The standard deviation of the averages for each particular sample size was then calculated and plotted. The dotted line in Figure 7.21 corresponds to these results. The solid line in Figure 7.21 represents the theoretical standard deviation according to (7.6).
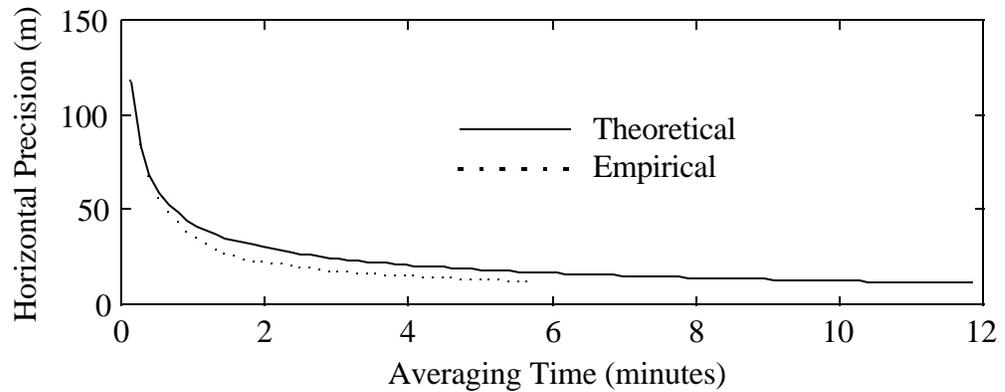
Figure 7.21   Effect of Averaging on Horizontal Precision
                   (using 8-second intervals)

The results of Figure 7.21 indicate a six fold improvement in the horizontal precision when position fixes over 2 minutes are averaged. The horizontal precision drops

177

from 118 m to approximately 20 m. The 425 position fixes obtained span approximately 56 minutes and the average interval between fixes is then 7.9 s. Therefore, a 2-minute averaging time represents 15 fixes. Averaging for longer than 3 minutes does not significantly improve the horizontal precision. Therefore, for the conditions of this test, a cellular 911 call need last for less than 3 minutes in order to derive maximum benefit from the averaging. With a shorter time interval between fixes, the averaging time necessary to obtain a 20 m horizontal precision will be even less.

7.4.4.5 Required Initial Position Accuracy

It was demonstrated in Chapter 6 that the accuracy of the initial position provided to LS affects the ability of LS to converge. Figure 7.7 suggests that the Plane Intersection method is suitable for providing the initial position. To confirm this for the present data set, one particular epoch was chosen and a grid of initial positions used to begin the LS process. The positional error for the epoch chosen was only a few metres. In Figure 7.22 each initial position is plotted by either a cross or circle. Those initial positions which result in the convergence of LS are marked by a cross; those that result in the divergence of LS are indicated by a circle. The four filled-in circles again represent the four Cellocate™ receiver sites.
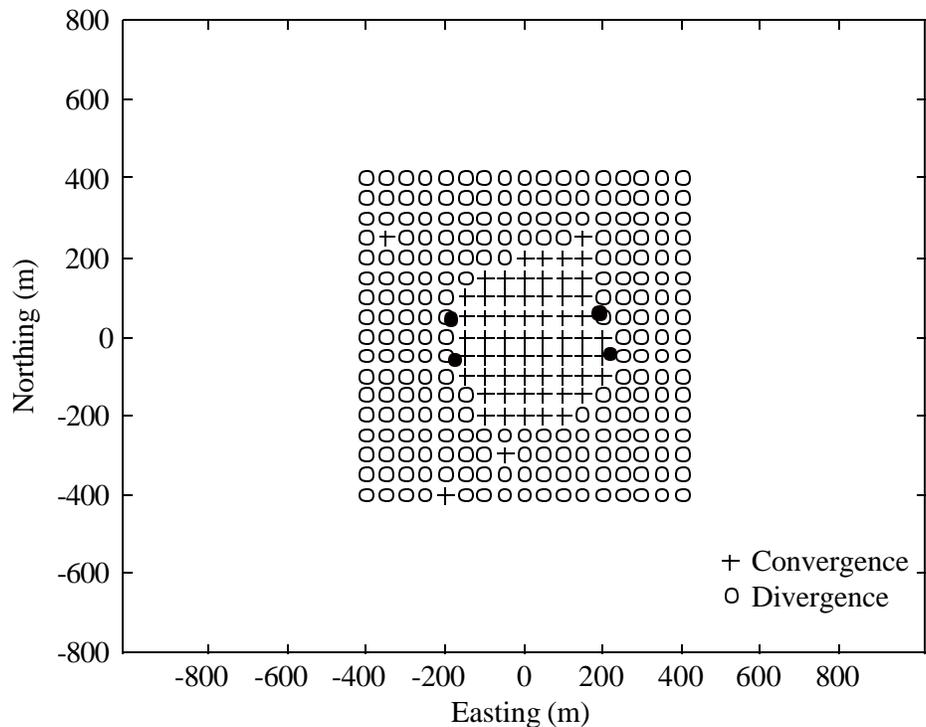
Figure 7.22   Results of Various Initial Positions for a Data Epoch at which
Least Squares is 5.6 m in Error

According to Figure 7.22, any initial position within approximately 200 m of the true transmitter position, results in the convergence of LS to the same solution. Comparing Figure 7.22 to Figure 7.7, one sees that a vast majority of the Plane Intersection solutions fall within the area of convergence. Less than 5% of the Plane Intersection solutions cause LS to diverge for the particular data epoch under consideration.

In a sense, the data epoch used in Figure 7.22 is a worst case scenario. Note that the true transmitter position, LS solution, and centroid of the polygon formed by the Cellocate™ sites are more or less equal. This results in a uniformly circular convergence region of minimum area. The same grid of initial positions was used with other data epochs. For those data epochs for which the LS solution was accurate (say within 50 m), the area of

convergence was very similar to that of Figure 7.22. For cases in which the LS solution was very much in error, the convergence area looked quite different. Consider the case in Figure 7.23. Here the LS solution is 449 m in error and indicated by the *. Note that the convergence area is now much bigger but still includes the circle of radius 200 m centered at the transmitter's true position.
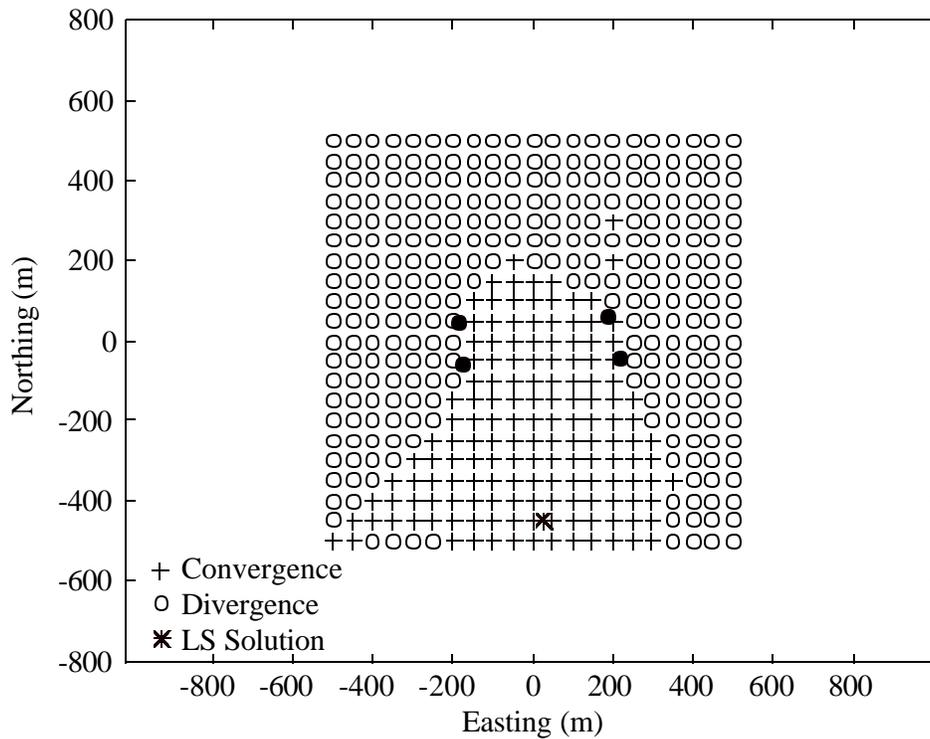
Figure 7.23   Results of Various Initial Positions for a Data Epoch at which   Least Squares is 449 m in Error

To conclude, the Plane Intersection solution for each data epoch was used as the initial position for LS. The results are almost identical to those of Figure 7.18 in which the approximate position supplied to LS was truth plus 100 m. The number of epochs for which LS converged dropped from 425 to 402. The DRMS, horizontal precision, 67th percentile,

and 95th percentile all dropped by 2 to 3 m. Therefore, for the data set analyzed in this section, Plane Intersection is capable of providing an initial position accuracy sufficient for LS to converge to a solution if it will in fact converge at all. However, it should be remembered that this is only true if lowpass filtered bias differences are used.

CHAPTER 8

OPERATIONAL FIELD TESTS

8.1 Telus Mobility Cell Sites

In order to test Cellocate™ in a city environment and with actual cell sites, an agreement was reached with Telus Mobility to install the Cellocate™ system in four Telus cell sites in Calgary. Cell sites in the northeast quadrant of the city were chosen since both the Cell-Loc Inc. and Telus Mobility offices are located in that area. Figure 8.1 is a plan view of the area. The black dots show the locations of the cell sites; the major thoroughfares are indicated by lines; shaded areas represent the approximate propagation environments; and the transmitter locations are marked by asterisks. Also shown is the direction of the main antenna lobe at each cell site.
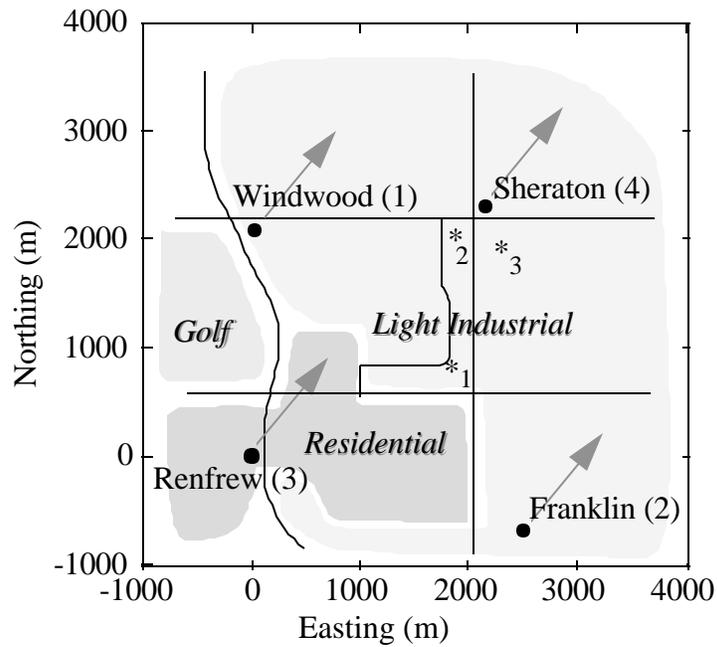
Figure 8.1   Telus Mobility Test Area

The four cell sites approximately form a box 2 km on a side. The Windwood cell sites sits atop a six story office building, one of the two highest buildings in the immediate area. The Sheraton cell site antenna is located on the roof of the Sheraton Cavalier Hotel which is eight stories high. Both the Franklin and Renfrew cell sites are located in utility huts with the antennas mounted on towers. The Renfrew tower is approximately 25 metres high and the Franklin tower is about 50 metres tall. The Windwood, Sheraton, and Franklin sites are located in light industrial areas whereas the Renfrew site is located in a residential neighborhood. Three different propagation areas are indicated in Figure 8.1. The darkest area corresponds to residential and the lightest corresponds to light industrial. These two areas are, to a certain extent, comparable to the suburban and urban areas respectively, assumed in the simulations of Chapters 5 and 6. The light industrial area largely consists of buildings anywhere from one to six stories. There is a fair amount of open space in this area and consequently the density of the buildings in this area is not high. Therefore, it is just as likely for LOS propagation to exist as not. However, due to the abundance of reflecting surfaces, it is very likely that multipath always exists. In residential areas, multipath is most likely due to reflections from nearby objects and the multipath delay is therefore small.

Almost all of the Telus Mobility cell sites in the city are sectorized due to the use of directional antennas. The common visualization of a cell site occupying the geographical centre of the cell's coverage area is not the norm. In the downtown area, many of the cell site antennas are located atop tall buildings and are tilted down in one direction or another. In other areas the choice of convenient locations for the cell site is limited and it is not possible to position the cell site at the centre of the intended coverage area. The four cell sites used in these tests are directional. Due to strong coverage by a cell site atop the tallest building in the Calgary's downtown (to the southeast of the sites in Figure 8.1), the cell sites of Figure 8.1 are directional to the northeast.

This limited the choice of locations from which to transmit the cellular signal. The Sheraton and Franklin cell sites were the most problematic. The Sheraton antenna pointed away from the interior of the box formed by the four sites. However, any transmitter to the northeast of Sheraton and, therefore, within the main lobe of its antenna was too far distant

to be heard by the Renfrew site. If the transmitter was to the southwest of Sheraton it could by heard by Renfrew but was now in the rear lobe of the Sheraton antenna and unless very close could not be received. Therefore, a separate omni-directional antenna was used at the Sheraton cell site. The Franklin cell site was unable to receive from anywhere west of an easting of approximately 1500 m in Figure 8.1. Possible areas for a cellular transmitter were therefore limited to the right half of the box formed by the cell sites. Any location to the east or north of this box was too distant from Renfrew. Any location south of the box was outside of Windwood's main antenna lobe and any point west of the box's centre was beyond the reach of Franklin.

Three locations from which to transmit the cellular signal where chosen. They are indicated by the asterisks in Figure 8.1 and labeled as 1, 2, and 3. A transmitter at these three locations could be consistently heard by the four cell sites. Location 1 was on the side of the road shown nearby in Figure 8.1 and LOS to Sheraton and Franklin. It was certainly not LOS to Renfrew and probably not LOS to Windwood. The area surrounding location 1 was somewhat open with the exception of a few one or two story buildings to the east and north and a multi-story hotel approximately 250 metres to the southwest. Location 2 was located on an access road to a parking lot. The immediate area was open. Approximately 50 metres to the west were several two or three story buildings in high density. This prevented LOS propagation to both the Windwood and Renfrew sites. Sheraton was very near to the northeast and obviously LOS. The Franklin site was also LOS. The third location was located in the parking lot of the Travel Lodge Hotel (approximately six stories). The hotel itself was just to the southeast of this location. The area to the west of this location was densely developed with an abundance of reflecting objects. Therefore, Renfrew and Windwood were not LOS whereas Franklin and Sheraton were.

The coordinates of the four cell sites were surveyed using DGPS. The reference station used was a point on the Cell-Loc building roof which in turn was tied to one of the control pillars atop the Engineering building at the University of Calgary. The baseline from the Engineering building to Cell-Loc is 8 km. Semikin™ was used to process this GPS baseline data which spanned approximately 1.5 hours. The accuracy of the double

differenced, fixed ambiguity coordinates of the point on the Cell-Loc rooftop is better than 3 cm. This point was then used as the reference station for surveying the cell sites and transmitter locations. Approximately 40 minutes of GPS data was collected at each cell site and the baseline from Cell-Loc to each cell site was 3 km or less. Semikin™ was again used and the double differenced, fixed ambiguity results should again be accurate to within 3 cm. The latitude and longitude of each cell site was then converted to Universal Transverse Mercator (UTM) northing and easting for use in the positioning software.

The coordinates of the transmitter locations were also obtained with DGPS. During testing, a GPS receiver was installed on the roof of Cell-Loc to act as the reference. Another GPS receiver was mounted in the test automobile which also contained the cellular transmitter. As a result a DGPS position for the cellular transmitter was always available. The positions of the three transmitter locations were obtained from a semi-kinematic survey. The occupation time at each of the points was between 5 and 8 minutes. The accuracy of the transmitter location coordinates is expected to be better than 5 cm, more than an order of magnitude better than required by this application.

The FCC specification of 125 m (67%) is for horizontal position. Therefore, height was not estimated in the simulations of Chapter 6 or the static multipath test of Chapter 7. Indeed, height differences in the static multipath test were altogether neglected. For the purpose of positioning, height differences are also neglected in this chapter. The measured DGPS coordinates of the cell sites and transmitter locations are three dimensional. In section 8.4.1 three dimensional ranges from the cell sites to the transmitter locations are subtracted from TOA differences. For those results, height is obviously accounted for. However, when horizontal position is estimated, height differences are disregarded since they are insignificant compared to the horizontal distances involved. The largest height difference is 61 m and occurs between the Franklin site and location 2. However, 61 m is only 2.1% of the three dimensional range of 2876.8 m. This corresponds to a difference in range of 60 cm when this height difference is neglected. The largest height difference in terms of percentage of the range is 24 m between the Sheraton site and location 2. The 24 m is 7.5% of the three dimensional range of 322.3 m. Neglecting this height difference

introduces an error of 90 cm. These errors are insignificant compared to TOA estimation error and multipath. However, in areas such as downtown, where cell sites often sit atop tall office buildings, it is quite possible that height differences will make up the majority of the propagation distances. In that case height must be accounted for. Estimating height, however, will reduce the degrees of freedom by one and may be difficult due to poor vertical geometry. If the participating cell sites all have similar elevations, the vertical DOP (VDOP) will be large.

8.2 Test Equipment

A Cellocate™ receiver and GPS receiver were mounted in each of the four Telus cell sites. The Cellocate™ receiver numbers are indicated in Figure 8.1 next to each site name. As discussed previously, at all sites with the exception of Sheraton, the actual cell site antenna was used. The cellular signal provided to the Cellocate™ receiver, was tapped from the cell site multicoupler. The GPS receiver was fixed to the precisely surveyed cell site coordinates and thereby operated in time transfer mode. Before the tests, each of the four Cellocate™ receivers was turned on and allowed to reach a stable operating temperature. From the bias test to follow, it was determined that approximately 20 minutes was required for this purpose. Once all the receivers were running for at least 20 minutes, the data logging software was activated and data collection commenced. The same transmitter used in the static multipath test was used here. The antenna was mounted on the automobile roof and the cellular precursor was transmitted every 2 seconds.

8.3 Cellocate Ô Biases

8.3.1 Cellocate Ô Receiver Biases

The bias of each Cellocate™ receiver was first analyzed in the laboratory to determine its stability. The transmitter was connected to each receiver in turn via a cable. Both TOA at the receiver and TOT at the transmitter were recorded for each transmission. The TOT was obtained by pulsing a second GPS receiver which shared the same antenna as the GPS receiver which recorded the TOA. Both the GPS receivers were operated in time

transfer mode. The transmission rate was once every 2 seconds and data was collected for approximately 30 minutes.

The TOT at each epoch was subtracted from the corresponding TOA to produce the bias plots of Figure 8.2. Note that the biases are expressed in metres as opposed to seconds and that the vertical scales are relative; for each plot the first TOA-TOT difference was subtracted from all others. Also included in each plot is the standard deviation, expressed in metres, of the last half of the data. It is obvious that each receiver's bias settles to some stable value after a time period of 10 to 15 minutes. As discussed in the previous chapter, these bias changes are temperature dependent. In this particular test, each receiver was not allowed to warm up prior to data collection. Therefore, what we see in Figure 8.2 is the warm up period of each receiver. The trends in Figure 8.2 could not be due to the transmitter since it was running continuously.
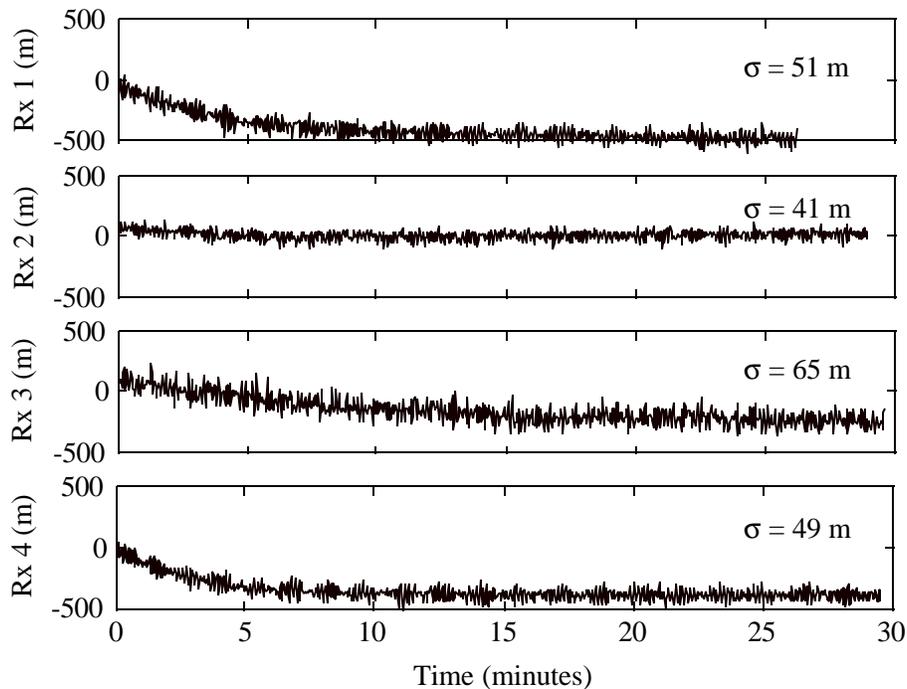
Figure 8.2   Cellocate™ Receiver Biases Measured in Lab

Upon reaching steady state, the standard deviation in the biases ranged from 41 m to 65 m. This is slightly better than the raw-filtered bias difference standard deviations of Table 7.1 because of the use of shielded cable as the propagation medium here. For the biases of Figure 8.2, the standard deviation of the bias differences will on average be 52 m * $\sqrt{2}$ = 73.5 m. The average bias difference standard deviation of Table 7.1 is 79.5 m. Figure 8.2 agrees with Table 7.1 in terms of relative standard deviations. Both sets of results suggest that receivers 1 and 3 are noisier than receivers 2 and 4.

8.3.2 Cellocate Ô Site Biases

A different approach was taken to measure the site biases for the Telus Mobility tests than that used for the static multipath test. The biases are obviously different here than in the case of the static multipath test since they now include propagation delay due to the cell site equipment. Although it is possible to measure the bias differences by differencing TOAs and subtracting differences in range to a transmitter of known position, as pointed out earlier, the bias differences so derived will contain multipath effects unique to the transmitter position. It was desired to measure the biases in such a way as to ensure their independence of the transmitter position.

Therefore, the bias at each cell site was measured by locating the transmitter as close as possible to the cell site antenna. In this way, it was assumed that the measured propagation time consisted only of the propagation time through the cell site equipment, and the receiver bias. No multipath effects should be included in these measured biases. As in the lab tests, the TOA at the receiver was recorded as well as the TOT at the transmitter. The difference between the two time tags, assuming the propagation distance between the antennas is zero, is the site bias. Because the TOT recorded was not the exact time at which the signal left the antenna, a bias exists in the measured site bias. This bias was removed by differencing between the measured site biases. The TOT bias was identical for all sites and therefore cancels out when bias differences are formed.

The distance between the transmitter antenna and receiver antenna ranged from approximately 10 m at the Windwood and Sheraton sites to the height of the towers at the

Renfrew and Franklin site. In any case the propagation distances were ignored since they are minor compared to the bias differences being measured. Expressed in terms of distance, the bias differences for the Franklin site as reference are on the order of 3 to 4 km. With the Renfrew site as reference the bias differences are anywhere from 1 to 4 km. In the worst case of the Franklin-Renfrew bias difference, neglecting the propagation distance results in a bias difference error of approximately 2%.
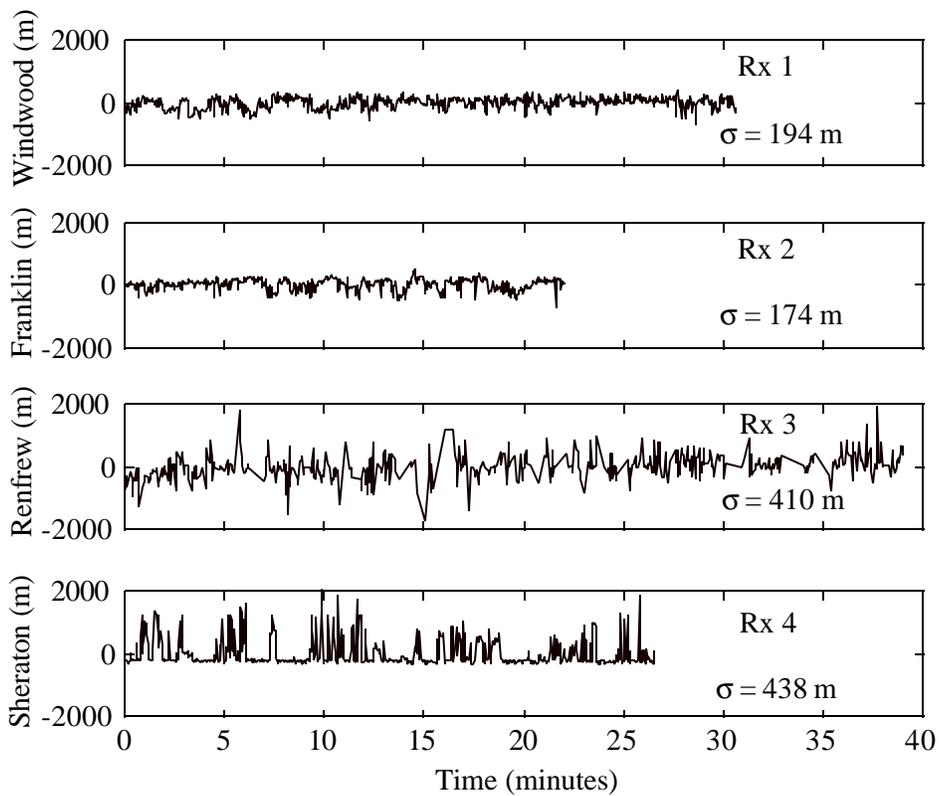
Figure 8.3   Measured Telus Mobility Cellocate™ Site Biases

The measured site biases are plotted in Figure 8.3. The biases plotted are the measured biases minus the corresponding mean. The transmission rate was again once every

189

2 seconds. The standard deviation of each plot and the receiver located at each site are also included. Note that the Cellocate™ receivers were allowed to warm up prior to data collection. The trends visible in Figure 8.2 are therefore absent from Figure 8.3. Compared to the bias differences of Figures 7.2 through 7.5, the bias differences here are slightly better behaved in terms of long term drift. However, in comparing the two cases it is important to remember that the time span of the data in Chapter 7 is almost one hour whereas the data of Figure 8.3 spans between 20 and 40 minutes. In any case, the marginal improvement is perhaps due to the fact that the Cellocate™ receivers in this case were located within the Telus Mobility cell sites in which the ambient temperature is strictly controlled. The transmitter was located outside and still subject to temperature variations.

   In terms of relative magnitude there is some correlation between the standard deviations of Figure 8.3 and Figure 8.2. Receivers 1 and 2 are very similar and the best behaved of the four receivers. Receiver 3 has a larger standard deviation than receivers 1 and 2 although by a larger ratio in Figure 8.3 than in Figure 8.2. In general, the standard deviations of the receivers are four to eight times larger than in the laboratory. This of course may be attributed to the propagation channel. In the laboratory the channel was a shielded cable which delivered a high SNR signal with no interference. In the field, the signal was transmitted over the air and therefore subject to greater attenuation loss, interference such as multipath, and noise from the cell site equipment. Recall that the simulations of Chapter 5 concluded that TOA estimation noise is a function of SNR.

   There is also a certain degree of correlation between Figure 8.3 and the results of Table 7.1. In section 7.4.2 it was postulated that a high standard deviation of any bias difference involving the NW site was due to the location of the NW site. The results here suggest that it is far more likely that the receiver at the NW site was to blame. Receiver number 3 was used at the NW site in Chapter 7 and at the Renfrew site here. It has a high standard deviation in both cases, possibly due to poor sensitivity. In addition, notwithstanding that the biases here and those of Chapter 7 were all measured over the air, the biases here have larger standard deviations even though the propagation distances are

190

much shorter. This may be due to white noise in the measured TOT. No TOTs were used in Chapter 7.

The bias of receiver 4 is rather odd when compared to the rest. In the laboratory, there was little difference between receiver 4 and receivers 1 and 2. In the field, the noise in the bias of receiver 4 can hardly be called white. It instead seems to have two characteristics. Throughout the some 27 minutes of data, lies a bias similar to that of receiver 2 in terms of standard deviation. At intervals throughout the data the bias jumps in a positive direction by up to 2000 m. Because these jumps are in a positive direction, one might suspect multipath. Receiver 4 was located on the roof of the Sheraton Cavalier Hotel. Structures on the roof where certainly capable of generating multipath but not on the order of 500 m to 2000 m as Figure 8.3 suggests. If multipath is to blame, the source must be a distant reflector. Other multistory buildings exist in the immediate area but are not as tall as the Sheraton. In any case, the bias for the Sheraton site will be a weak point in the positioning process.

The biases of Figure 8.3 were used in some of the positioning results to follow. Because these biases were measured on a separate day (Oct. 28, 1996) from the collection of other data, they cannot be applied on an epoch by epoch basis. Instead, average biases were calculated from the data of Figure 8.3 and used to correct TOA data. The average biases and the corresponding bias differences with the Franklin site (receiver 2) as reference are given in Table 8.1. Note that they are expressed in distance rather than time. For all positioning results to follow, the Franklin site is used as the reference site for the purpose of forming TOA differences.

Table 8.1   Average, Measured Biases and Bias Differences (Oct. 28, 1996)
                         with respect to Franklin

Cell Site	Measured, Average Bias (m)	Bias Difference (m)
Windwood (rx 1)	1 280 513	3261
Franklin (rx 2)	1 277 252	N/A
Renfrew (rx 3)	1 281 307	4055

Sheraton (rx 4)	1 280 022	2770

8.4 Static Data

Static data was collected on two separate occasions, Oct. 3 and Oct. 24, 1996. The equipment setup on these two days was identical to that of Oct. 28, the date the biases were measured. In this way the biases measured on Oct. 28 could be used with the data collected on the other two dates. On Oct. 3, 1996, data was collected for the transmitter at locations 1 and 2 in Figure 8.1. On Oct. 24, 1996, the transmitter visited all three locations. Each transmitter location was occupied for approximately 10 minutes. Only those epochs at which all four cell sites received a transmission were used for positioning. Sensitivity was poorest for the Renfrew site due to the lack of a LOS path and perhaps the use of receiver 3 there. Windwood was the next worse. Most often epochs were discarded because Renfrew or Windwood did not receive at that particular epoch. The number of common epochs for each transmitter location on each date of data collection is given in Table 8.2. Also included is the time span which these epochs cover. The time between common epochs ranged from 3.1 to 5.4 seconds. If all sites receive every transmission, the interval would be 2 seconds per epoch.

Table 8.2   Static Data Characteristics

Txer Location and Date	No. of Data Epochs	Time Span	Interval
Loc. 1 Oct. 3, 1996	103	5 min 21 secs	3.1 s/epoch
Loc. 1 Oct. 24, 1996	94	7 min 29 secs	4.8 s/epoch
Loc. 2 Oct. 3, 1996	129	7 min 7 secs	3.3 s/epoch
Loc. 2 Oct. 24, 1996	66	5 min 54 secs	5.4 s/epoch
Loc. 3 Oct. 24, 1996	97	5 min 29 secs	3.4 s/epoch

8.4.1 Bias Differences

8.4.1.1 Multipath Effects

As in Chapter 7, the first step in processing the data was to difference TOAs between the cell sites and subtract the corresponding known range to transmitter difference. This essentially removes the propagation distance from the data. What remains is the cell site bias differences as well as errors due to multipath and TOA estimation. Figures 8.4 through 8.6 show these differences for three of the five data sets. These figures are similar to Figures 7.2 to 7.5 in that they contain the raw TOA differences with range difference subtracted, as well as lowpass filtered versions. They also contain a dashed horizontal line at the values of the bias differences of Table 8.1. In each figure the three plots are labeled according to the sites and the receivers used at those sites. Each plot also contains the standard deviation of the raw data. The same normalized cutoff frequency was used for the lowpass filters here as in Chapter 7.
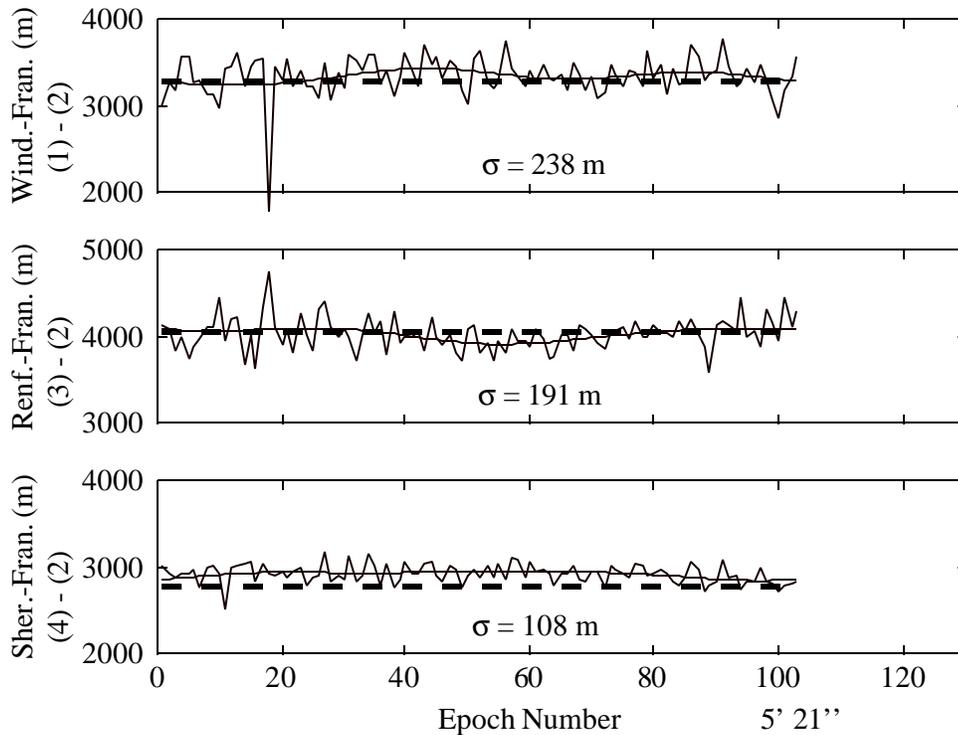
Figure 8.4   Raw and Lowpass Filtered Bias Differences (Franklin Reference)

Oct. 3, 1996 Loc 1 Data

        The difference between the lowpass filtered differences and the measured biases of Table 8.1 is an indicator of the multipath in the TOA difference measurements. Recall that the lowpass filtered differences contain the site biases, the mean TOA estimation error and multipath errors. The measured bias differences (dashed lines) should only consist of the site bias and the mean TOA estimation error. Multipath should be minimal, if it exists at all, since the transmit and receive antennas were placed as close as possible to one another when measuring the biases. Subtracting the measured bias differences from the lowpass filtered TOA differences, should eliminate the cell site biases from the TOA differences as well as the mean TOA estimation error.   What remains is multipath on the TOA difference measurement and any residual noise.
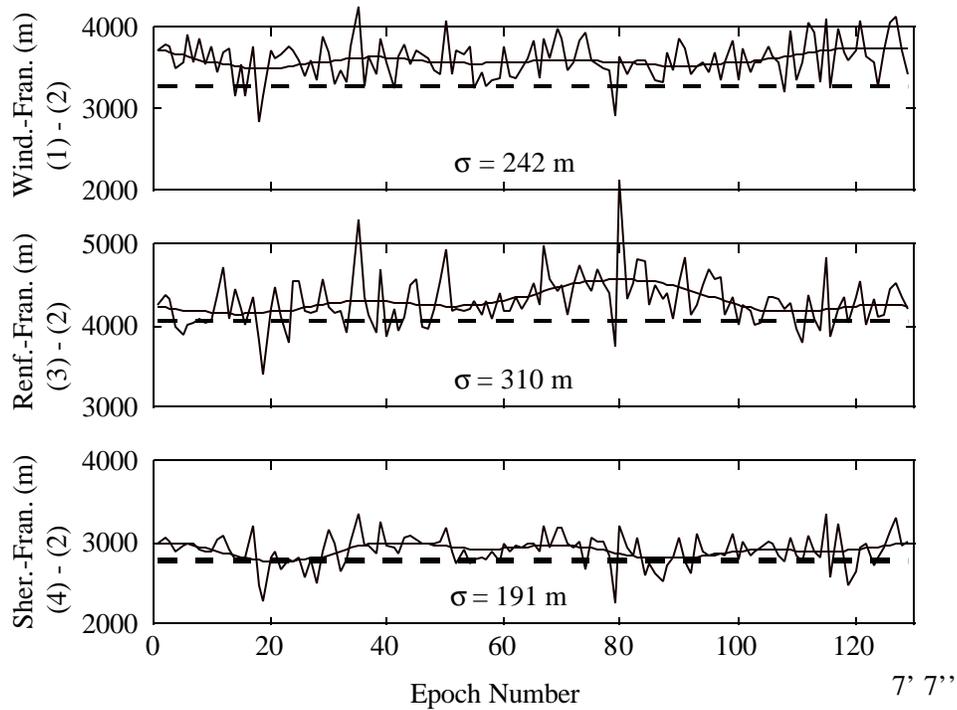
Figure 8.5   Raw and Lowpass Filtered Bias Differences (Franklin Reference)
                    Oct. 3, 1996 Loc 2 Data

The mean difference between the measured bias differences (dashed lines) and the lowpass filtered TOA differences was calculated for each cell site pair for each of the five data sets. These values are tabulated in Table 8.3. The cell sites are again identified both by the initial of the cell site name and the Cellocate™ receiver number. Note that any random noise still present will be removed when calculating the mean values of Table 8.3. Averaging the absolute values of all means for each particular transmitter location results in the following: Loc 1, 106 m; Loc 2, 182 m; Loc 3, 109 m.
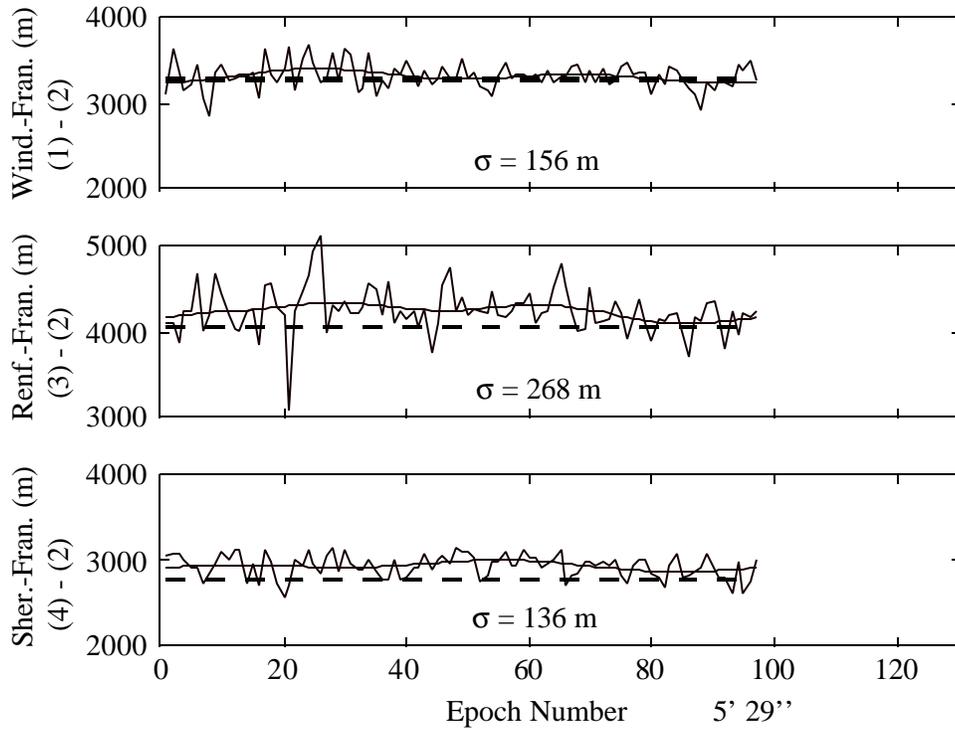
Figure 8.6   Raw and Lowpass Filtered Bias Differences (Franklin Reference)

Oct. 24, 1996 Loc 3 Data

Table 8.3   Mean Differences between TOA Differences (with Range

Differences   Removed)   and   Measured   Site   Bias

Differences

Data Set	W - F (m)	R - F (m)	S - F (m)	W - S (m)	R - S (m)	W - R (m)
	(1) - (2)	(3) - (2)	(4) - (2)	(1) - (4)	(3) - (4)	(1) - (3)
Oct. 3 Loc 1	71	-30	146	-76	-176	101
Oct. 24 Loc 1	-149	63	-38	-112	101	-213
Oct. 3 Loc	328	233	123	205	111	94

2						
Oct. 24 Loc 2	315	228	336	-21	-108	87
Oct. 24 Loc 3	50	181	151	-109	30	-131

When interpreting the results of Table 8.3, one must be careful to remember that these numbers are indicators of multipath on TOA differences. A number of small magnitude in Table 8.3 can indicate one of two different things; either there is little multipath in both of the TOAs involved, or both TOAs are corrupted by roughly the same amount of multipath. On the other hand, a number of large magnitude in Table 8.3 must indicate a significant multipath effect on at least one of the two TOAs involved. The sign of the number may indicate which of the TOA measurements is the leading contributor to a large magnitude. If one assumes that multipath always delays TOA, then a positive number implicates the numerus ex quo subductus while a negative number points to the numerus subducendus. One also needs to remember that the measured site bias differences are assumed to accurately represent the site bias differences only.

From Table 8.3, one may rank the sites, according to the assumptions made, in terms of the amount of multipath. This is done in Table 8.4. For each data set the sites are listed from that with the highest multipath effect to that with the lowest. Recorded between each adjacent pair of sites is the difference in multipath effect. The last row of the table contains the mean of the cumulative sum of the numbers for each column. For example, in the first column the mean of 103 m is the mean of 30 m, 30 m + 71 m, and 30 m + 71 m + 76 m.

Table 8.4   Relative Multipath Effects

Oct. 3 Loc 1	Oct. 24 Loc 1	Oct. 3 Loc 2	Oct. 24 Loc 2	Oct. 24 Loc 3
Sheraton	Renfrew	Windwood	Sheraton	Renfrew
76 m	63 m	94 m	21 m	30 m
Windwood	Franklin	Renfrew	Windwood	Sheraton

197

71 m	38 m	111 m	87 m	109 m
Franklin	Sheraton	Sheraton	Renfrew	Windwood
30 m	112 m	123 m	228 m	50 m
Renfrew	Windwood	Franklin	Franklin	Franklin
Mean = 103 m	Mean = 158 m	Mean = 228 m	Mean = 293 m	Mean = 133 m

From Tables 8.3 and 8.4 one may make the following observations. In general, multipath seems to be more significant for Loc 2 than for Loc 1 and Loc 3. Comparing the two dates for Loc 1 there is little consistency in terms of sign and magnitude. In fact, there is almost a complete reversal of the cell sites in the terms of the severity of multipath. Loc 2 is more consistent between the two dates. For both dates all sites experienced more multipath than Franklin. That Windwood and Renfrew experienced large amounts of multipath relative to Franklin is consistent with the known propagation environment. As discussed in 8.1, both Windwood and Renfrew did not have LOS paths to Loc 2. Sheraton, however did have LOS propagation with Loc 2 and it is therefore unclear why it would experience such a large multipath effect with respect to Franklin. It is quite possible that since the actual range between Loc 2 and Sheraton is only 322 m, reflections from the buildings due west of Loc 2 may have had amplitudes of the same order as the LOS ray. The Loc 3 results indicate large multipath for Renfrew and Sheraton and smaller multipath for Windwood and Franklin. One would expect small multipath effects for Franklin since it is LOS, but not for Windwood. Large multipath was expected for Renfrew but not Sheraton. It is not immediately obvious why Sheraton would experience large amounts of multipath for this location.

If the assumption is made that the lowest cell site in each column of Table 8.4 experienced no multipath, then the multipath effects in the other three cell sites ranges from 30 m to 336 m with the mean multipath effect ranging from 103 m to 293 m. This is entirely consistent with the Turin based multipath models used in the positioning simulations of Chapter 6. Compared to the urban probability density function illustrated in Figure 6.2, the results here may seem somewhat pessimistic. However, the PDF of Figure 6.2 is for the first

arrival only whereas the results reported here are indicative of the total multipath effect - both the delay of the first arrival and the effect of later arrivals. Secondary signal arrivals will obviously bias the TOA estimate in a positive direction.

8.4.1.2 TOA Measurement Noise

Although the data collected is only of short duration it is still instructive to observe the standard deviation of the raw, filtered, and raw minus filtered bias differences. These standard deviations, for all of the data collected, appear in Table 8.5.

Table 8.5   Standard Deviations of Bias Differences

		S_{raw} (m)	$S_{filtered}$ (m)	$S_{raw-filtered}$ (m)
	W(1) - F(2)	238	55	231
Oct. 3 Loc 1	R(3) - F(2)	191	60	182
	S(4) - F(2)	108	32	103
	W(1) - F(2)	250	71	239
Oct. 24 Loc 1	R(3) - F(2)	332	98	314
	S(4) - F(2)	158	61	145
	W(1) - F(2)	242	69	231
Oct. 3 Loc 2	R(3) - F(2)	310	125	283
	S(4) - F(2)	191	63	172
	W(1) - F(2)	183	83	161
Oct. 24 Loc 2	R(3) - F(2)	371	62	361
	S(4) - F(2)	151	65	134
	W(1) - F(2)	156	47	148
Oct. 24 Loc 3	R(3) - F(2)	268	79	254
	S(4) - F(2)	136	36	131

The standard deviations of Table 8.5 are significantly higher than those of Table 7.1. This is most likely due to longer propagation distances and lack of LOS for the Telus Mobility data. Of the Windwood, Renfrew, and Sheraton cell sites, the Sheraton cell site is

the only one for which it was certain a LOS path to the three transmitter locations existed. It was also the closest site, of the three just mentioned, to the transmitter locations. It would therefore be quite reasonable to assume that the received signal SNR at the Sheraton site was significantly higher than at the Windwood and Renfrew sites. The SNR at the Franklin site is also expected to be high since a LOS path exists from Franklin to the three transmitter locations.

The difference in SNR and receiver sensitivity are expected to be the major causes of the difference in standard deviation between the three cell site pairs of Table 8.5. For all but the Oct. 3 Loc 1 data set, the Renfrew-Franklin bias difference has the highest standard deviation whereas the Sheraton-Franklin bias difference has the lowest. This is consistent with the expected received SNR at the Sheraton, Windwood, and Renfrew sites and the poorer performance of Cellocate™ receiver number 3. Recall that receiver 3 was involved in the highest noise standard deviations in the static multipath test as well.

Comparing the two Oct. 3 data sets, there is little difference, as far as the Windwood site is concerned, between transmitter locations 1 and 2. The Renfrew-Franklin bias difference standard deviation for transmitter location 2 is approximately 100 m higher than for location 1. Again one would assume that a significantly longer propagation path and therefore lower SNR would cause higher TOA estimation noise. For the Oct. 24 data sets there is little difference between Loc 1 and Loc 2 with respect to the Sheraton-Franklin bias difference whereas for the Oct. 3 data sets there is a 70 m difference. In making all of these comparisons one must recognize that the data collected only spans a few minutes. Standard deviations based on data spanning a longer time period may be more consistent with expectations.

The average raw-filtered standard deviation from Table 8.5 is 206 m. In the absence of any long term drift or bias in the measurements and an HDOP of 1, one would expect an instantaneous horizontal precision of approximately $206/\sqrt{2} * 1 = 146$ m.

8.4.2 Plane Intersection Results

Plane intersection was tested with the field test data. Again, its performance was dependent on the biases used. Using the bias differences of Table 8.1, the position results for the Oct. 3 Loc. 1 data were extremely poor. The results are similar to those of Figure 7.6, for which average site biases were employed in the static multipath test. As in that case, all four accuracy and precision measures here are in the thousands of kilometers.

Using filtered bias differences the five sets of data were processed with plane intersection. Figure 8.7 contains the position results for the Oct. 3 Loc 1 data set. Plane intersection position results for all five sets of data are given in Table 8.6.
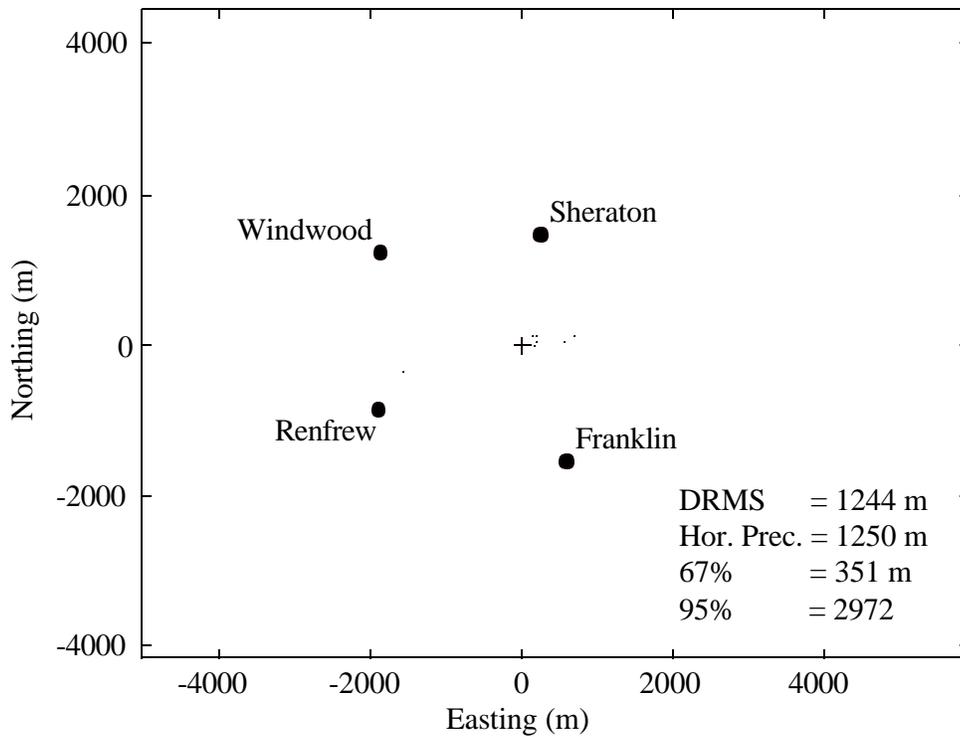

Figure 8.7   Oct. 3 Loc 1 Position Results Using Plane Intersection and
Lowpass Filtered Site Bias Differences

It is interesting to note that in Table 8.6, the 67th percentiles for the Loc 1 data are approximately one half of that of the other data sets. Because the bias difference standard deviations for Loc 1 are not lower than those for the Loc 2 and Loc 3 data sets another factor

must be causing the difference. This factor cannot be multipath since any constant multipath will be accounted for in the filtered bias differences used and any short term multipath would contribute to the raw minus filtered bias differences.

The only remaining factor is the sensitivity of plane intersection to geometry. The HDOP for Loc 1 and Loc 2 is 1.1 whereas the HDOP for Loc 3 is 1.5. This leads one to believe that there should be no difference in positioning accuracy due to geometry between Loc 1 and Loc 2 and an increase of 1.4 times in the position error of Loc 3 with respect to Loc 1 and Loc 2. However, this is only true for positioning algorithms which employ the design matrix **A** from which HDOP is derived. Plane intersection does not make use of the design matrix. Schmidt (Schmidt, 1972) investigated the sensitivity of plane intersection to geometry. He simulated networks of three or four stations. To one of the perfect TOA differences, corresponding to a specific transmitter point, he then added an error and obtained a position fix with plane intersection. This was done for a grid of transmitter points, the same error being added to the same TOA difference at all points. He then connected points with the same position error to produce error contour maps. Although none of Schmidt's networks are comparable to the network under consideration here, his simulations do demonstrate large position error for points close to one of the receiver sites, particularly for points outside the polygon formed by the sites. Outside the polygon there is also large position error for points close to the line formed by any two of the receiver sites. Both Loc 2 and Loc 3 fall into these areas of problem geometry.

Table 8.6   Position Results Using Plane Intersection and
                          Lowpass Filtered Site Bias Differences

	DRMS (m)	Hor. Prec. (m)	67% (m)	95% (m)
Oct. 3 Loc 1	1244	1250	351	2972
Oct. 24 Loc 1	4414	4362	399	2970
Oct. 3 Loc 2	1465	1461	643	2702
Oct. 24 Loc 2	6366	6414	982	4403
Oct. 24 Loc 3	3078	3049	700	2385

8.4.2.1 Required Accuracy of Initial Position

To determine if the plane intersection results of Table 8.6 are sufficiently accurate for use as initial positions for LS, a grid of initial positions was generated for each of the three transmitter locations. These initial positions were each used by LS for one particular epoch of data which results in a very small position error. For the Oct. 3, Loc 1 data set a data epoch was chosen for which the LS position error is 25 m. The results are given in Figure 8.8. Initial positions resulting in convergence are marked by a + whereas those resulting in divergence are marked by a o. The cell sites are indicated by the large black dots. The plane intersection results of Figure 8.7, represented by small black dots, are also shown.
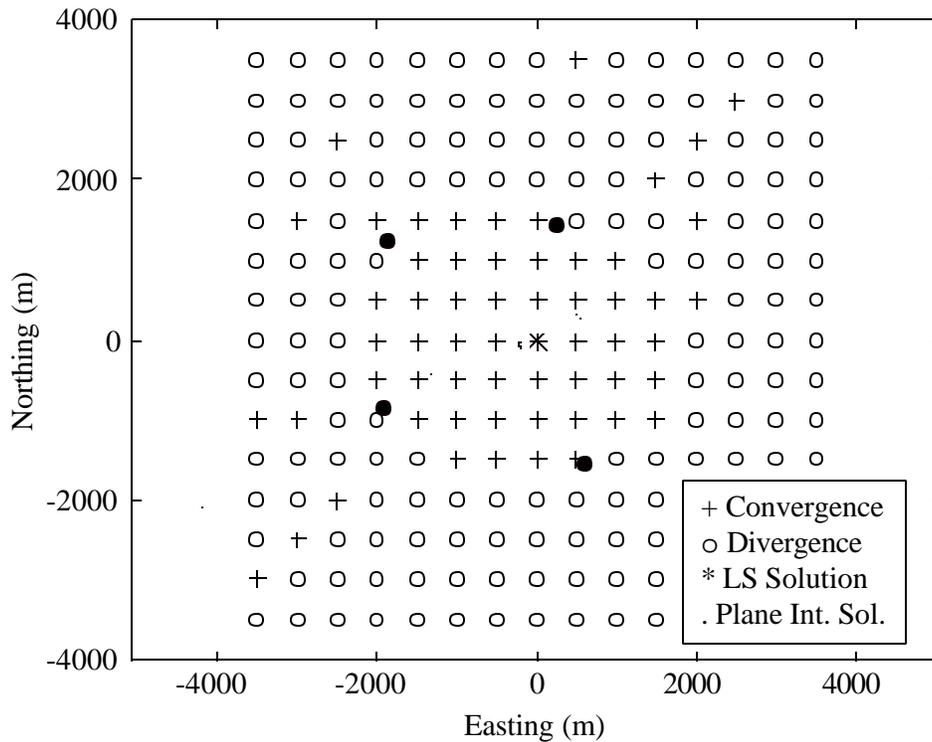
Figure 8.8   Results of Various Initial Positions for a Data Epoch at which LS is 25 m in Error (Oct. 3, Loc 1)

It is apparent from Figure 8.8 that the overwhelming majority of the plane intersection position fixes would result in convergence of LS for the particular data epoch under consideration. The figure indicates that the initial position must be within 1500 m to 2000 m of truth in order for LS to converge to the correct (25 m in error) position. Of the plane intersection fixes, 92% are within 2000 m of the true transmitter position. This particular data epoch, which results in a LS position error of only 25 m, is considered a worst case scenario in that the area of convergence is most likely the smallest for the geometry given. Recall Figure 7.23 which shows a larger convergence area when the LS position fix error is larger.
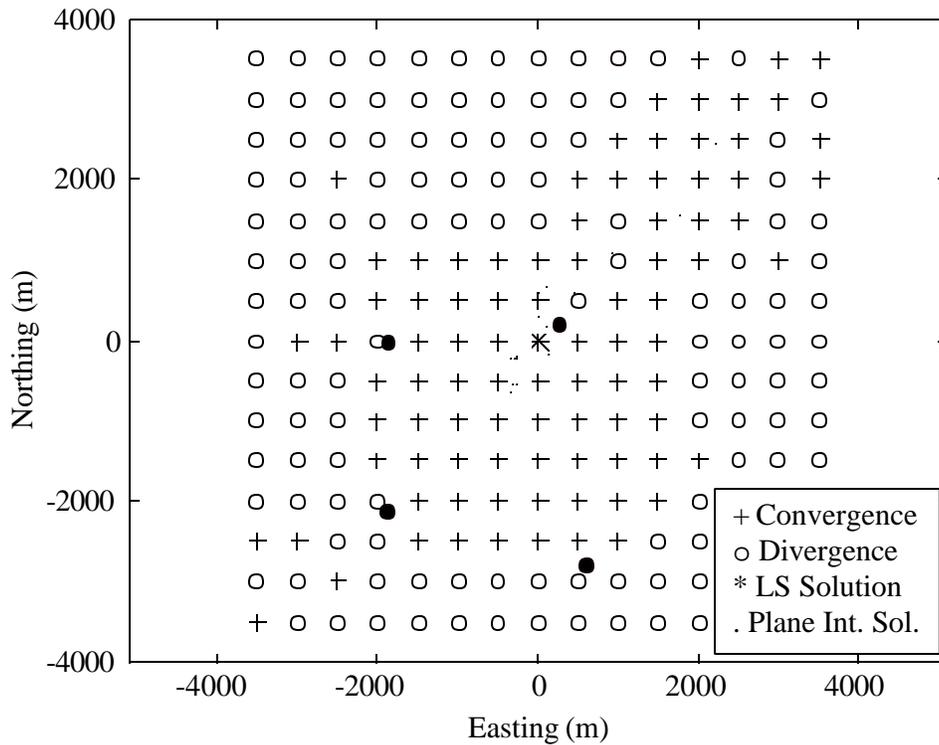
Figure 8.9   Results of Various Initial Positions for a Data Epoch at which LS is 58 m in Error (Oct. 3, Loc 2)

Figure 8.9 shows the LS convergence area for a particular data point from the Oct. 3, Loc 2 data set. This case may also be considered a worst case since the LS position fix is only 58 m in error. Note that in contrast to the circular convergence area of Figure 8.8, here the convergence area is a circle centered on the cell site network with an additional lobe of convergence to the NE. The plane intersection results are also shown and some of them fall in this lobe of convergence. Overall, 90% of the plane intersection fixes, when used as the initial position, result in convergence for this particular data epoch.

A LS convergence plot was also done for Loc 3. The data epoch chosen results is a LS position error of 28 m. Figure 8.10 illustrates the convergence area for this one data epoch and the plane intersection results for the entire data set. Note the familiar circular convergence area with an additional region of convergence to the east and northeast. Using the plane intersection fixes as initial positions for this one particular data epoch results in a convergence rate of 78%.
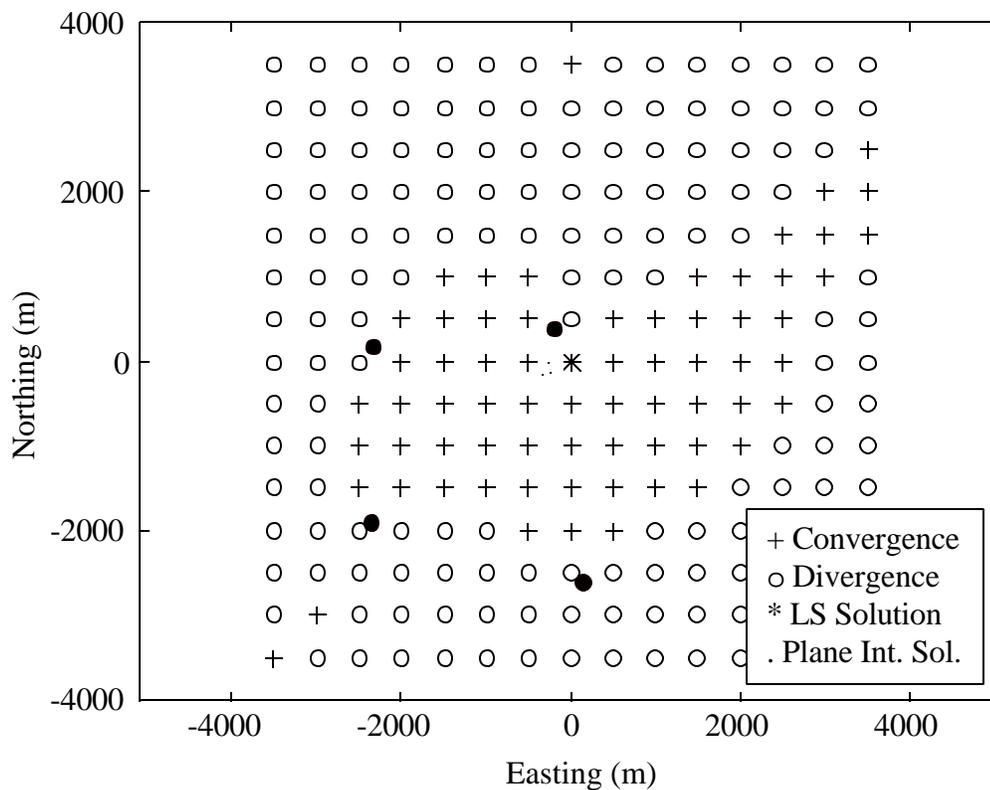
Figure 8.10   Results of Various Initial Positions for a Data Epoch at which LS

is 28 m in Error (Oct. 24, Loc 3)

For the three data sets considered here, plane intersection is capable of generating an initial position resulting in the convergence of LS in 78% to 92% of all cases. This is considered a pessimistic result since the data epoch used for each data set resulted in a very accurate LS position location. However, in order for plane intersection to achieve this level of accuracy, it was necessary to use lowpass filtered epoch by epoch cell site bias differences derived from the actual data and the known ranges to the transmitter. This is obviously not a practical solution since it requires a priori positional information. More work needs to be done in order to more accurately measure the cell site biases which may be used with any data and are not dependent on the transmitter's location. It is anticipated that with highly

accurate cell site biases and additional cell sites, plane intersection will be capable of providing a sufficiently accurate initial position to begin the LS iterative process. To separate the effect of plane intersection accuracy from LS accuracy in the results to follow, the initial position provided to LS will always be a point 100 m to the north and 100 m to the east of the true transmitter position.

8.4.3 Least Squares Position Results
8.4.3.1 Position Results

The five data sets were processed using root MUSIC to estimate TOA and LS to estimate position. Circular trilateration with TOA differences was again used in order to obtain residuals for the individual TOA observations. The Franklin site was used as the reference site. In order to make the results as realistic as possible, raw TOA differences were corrected with the average measured bias differences of Table 8.1. Lowpass filtered, epoch by epoch bias differences were not used. To follow are figures describing the results for the following three data sets: Oct. 3 Loc 1, Oct. 3 Loc 2, Oct. 24 Loc 3. Figures 8.11, 8.12, and 8.13 illustrate the position results, individual parameter estimates and residuals for the Oct. 3 Loc 1 data set. Figures 8.14, 8.15, 8.16 and 8.17, 8.18, 8.19 show the same for the Oct. 3 Loc 2 and Oct. 24 Loc 3 data sets respectively. As before, the small dots in the position results are the epoch by epoch position fixes and the + marks the location of the transmitter as determined by DGPS.
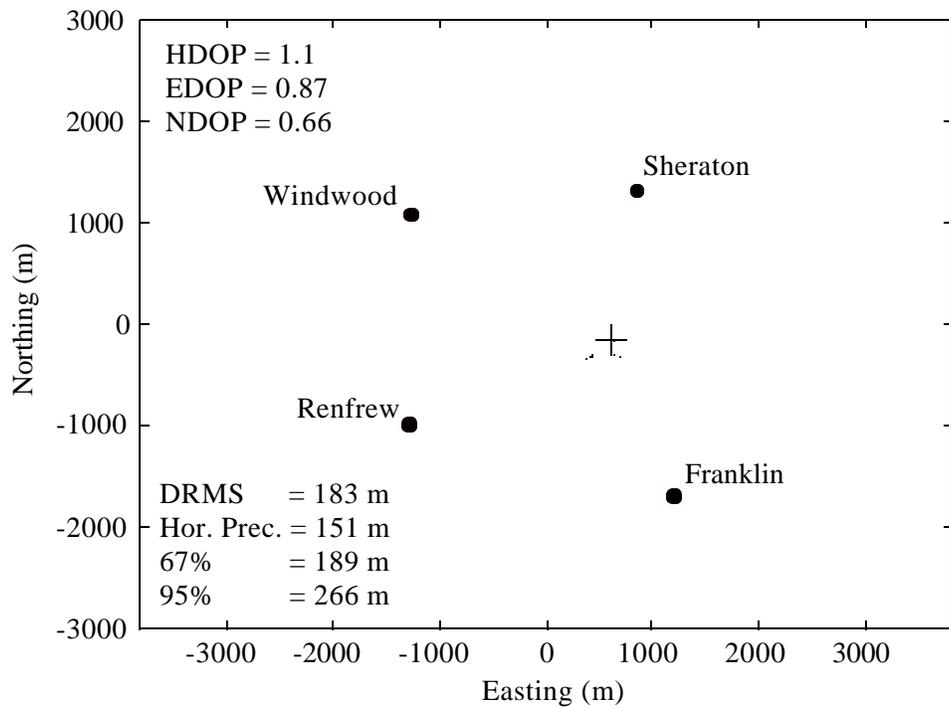
Figure 8.11   Position Results for Oct. 3 Loc 1
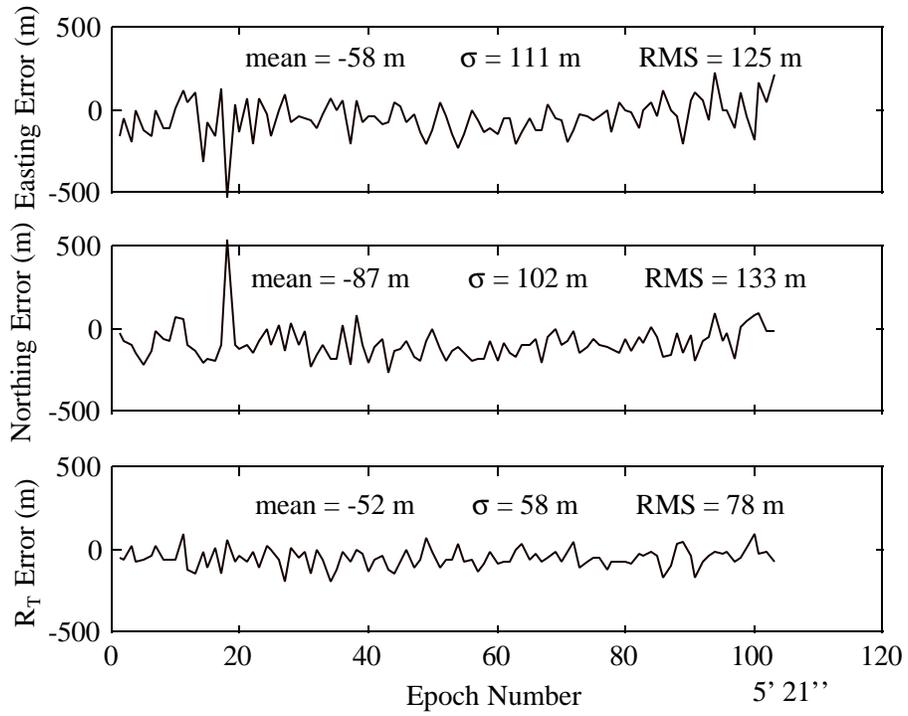
Figure 8.12   Parameter Estimates for Oct. 3 Loc 1
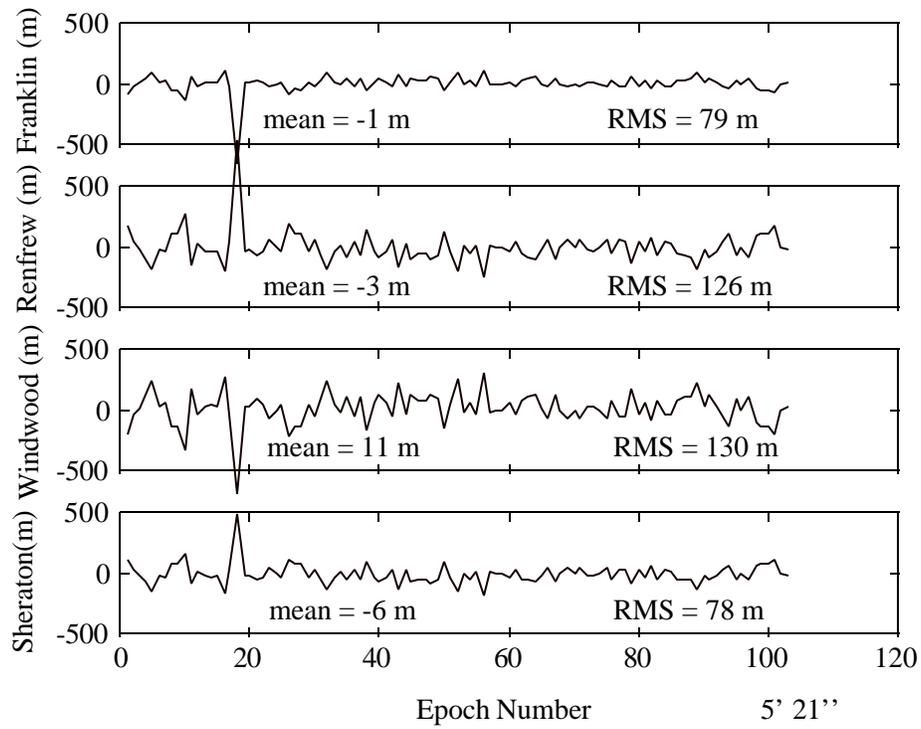
Figure 8.13   Observation Residuals for Oct. 3 Loc 1
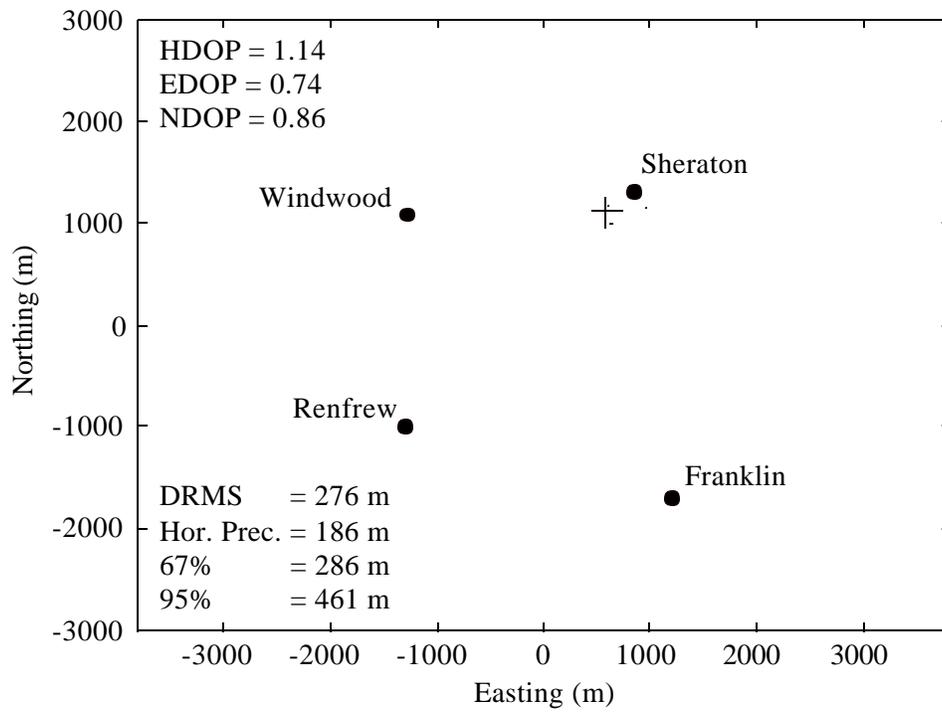
Figure 8.14   Position Results for Oct. 3 Loc 2
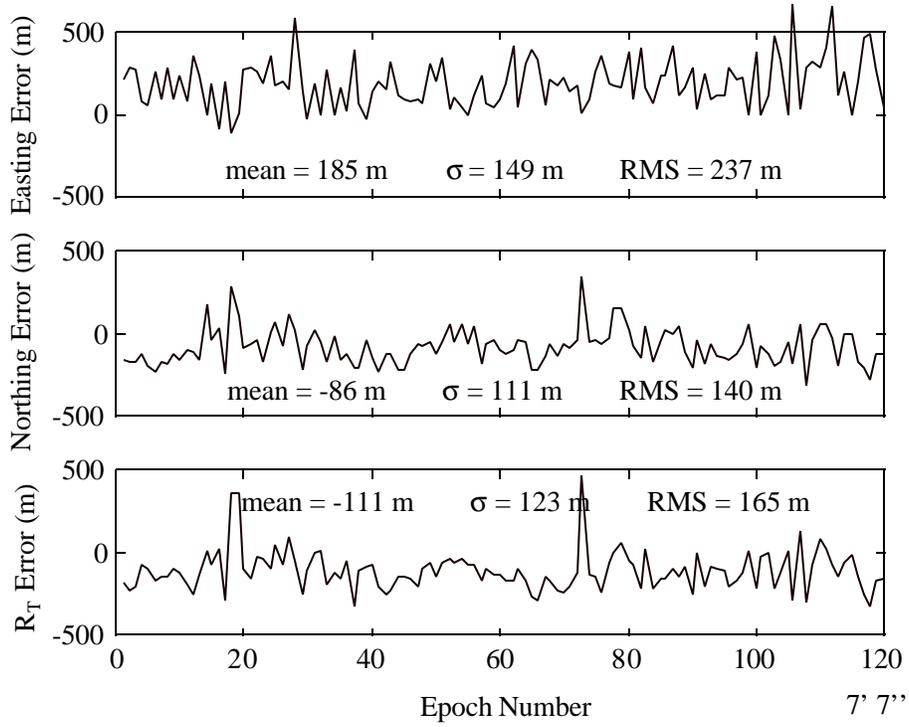
Figure 8.15   Parameter Estimates for Oct. 3 Loc 2
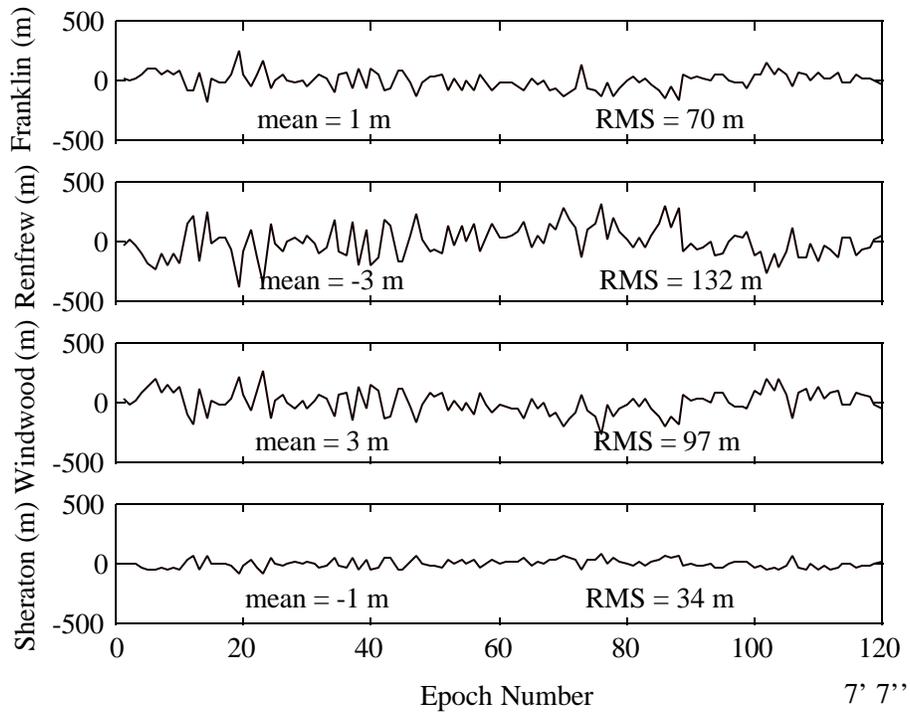
Figure 8.16   Observation Residuals for Oct. 3 Loc 2
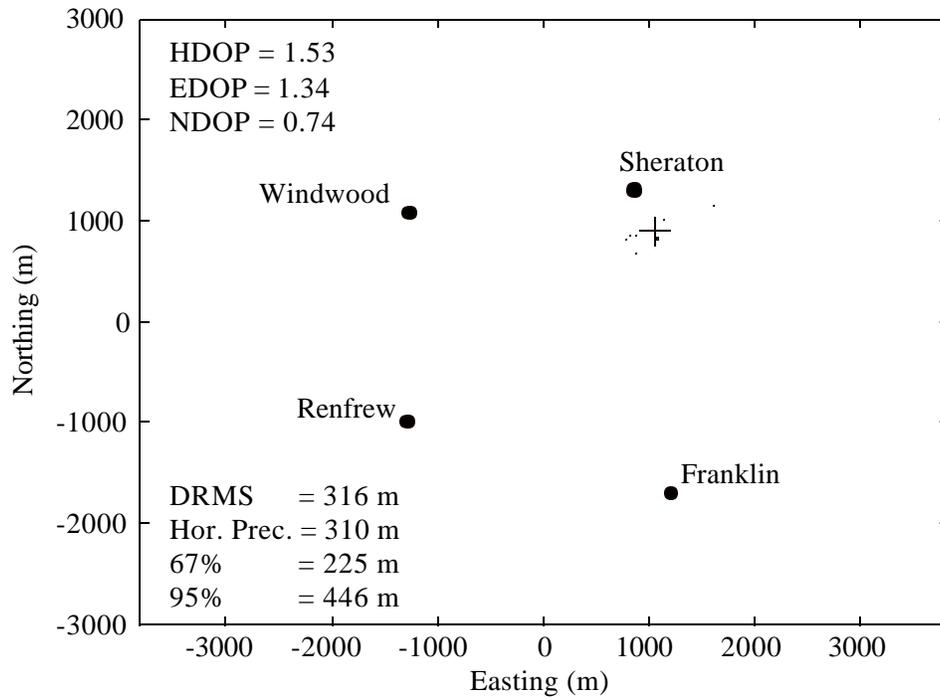
Figure 8.17   Position Results for Oct. 24 Loc 3
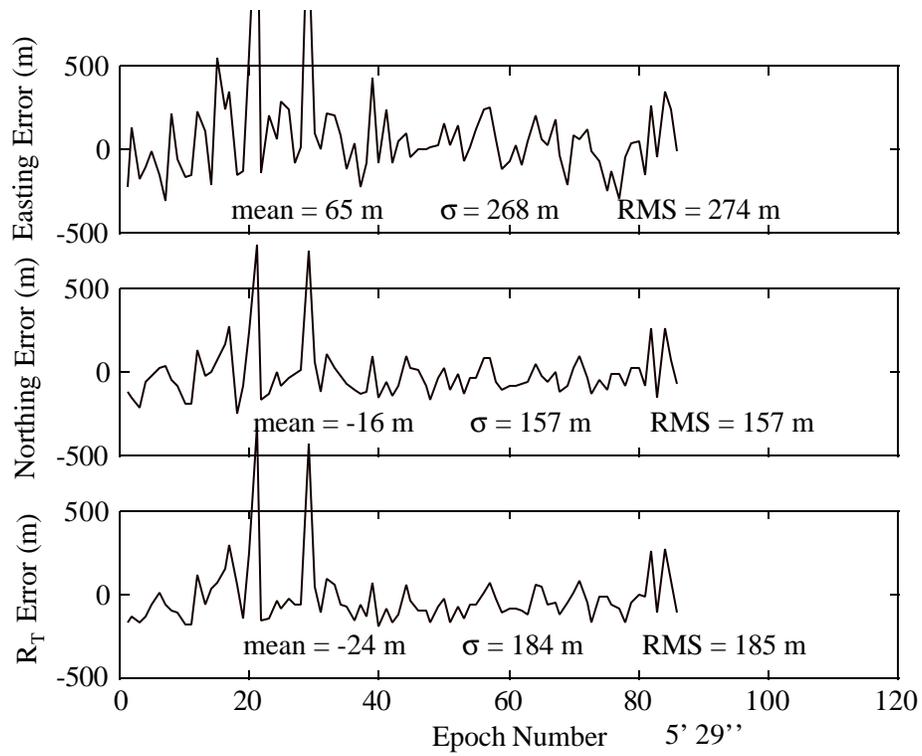
Figure 8.18   Parameter Estimates for Oct. 24 Loc 3
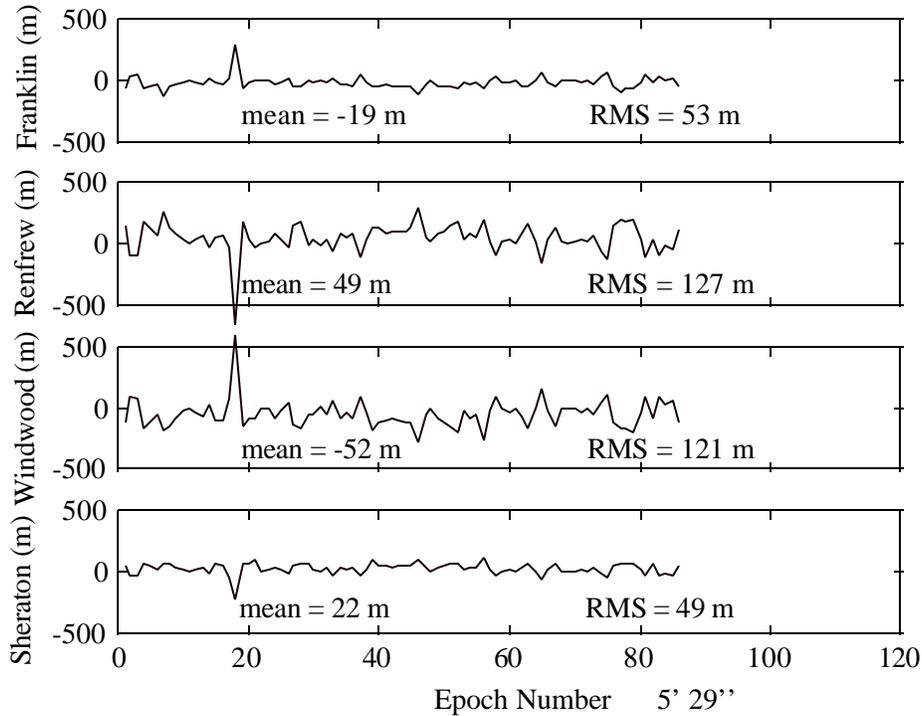
Figure 8.19   Observation Residuals for Oct. 24 Loc 3

The results for all five data sets are summarized in Tables 8.7 through 8.11. The horizontal position performance measures are given in Table 8.7 whereas individual parameter statistics are given in Table 8.8. The observation residuals appear in Table 8.9, their correlation coefficients in Table 8.10, and the observation redundancy numbers in Table 8.11.

It is clear from Table 8.8 that the position estimates are biased. This is also reflected in Table 8.7. For zero mean Gaussian errors, the DRMS and horizontal precision should be equal. For the Oct. 24 Loc 1 data set the two measures are only 7 m apart whereas for the Oct. 3 Loc 2 data the two numbers differ by 100 m. These position biases are due to the difference between the average measured biases and the lowpass filtered biases. As discussed in section 8.4.1.1, it is reasonable to assume that the differences in the measured

216

and filtered biases are due to multipath. However, since the same bias differences were used for all data, the fact that the magnitude of the biases in Table 8.8 vary from one date to the next, for the same location, suggests a change in the multipath environment. On the other hand, perhaps more data is required to say with certainty that the position bias does change significantly from day to day. If the position estimates were Gaussian, the DRMS and 67th percentiles would also agree. This is true here with the exception of the Oct. 24 Loc 3 data. In that case, outliers clearly visible in Figure 8.17, inflate the DRMS value.

Table 8.7   Position Results Using LS and Average Site Bias

Differences

	DRMS (m)	Hor. Prec. (m)	67% (m)	95% (m)
Oct. 3 Loc 1	183	151	189	266
Oct. 24 Loc 1	220	213	208	434
Oct. 3 Loc 2	276	186	286	461
Oct. 24 Loc 2	279	172	287	436
Oct. 24 Loc 3	316	310	225	446

Table 8.8   Parameter Estimate Error Statistics

	Easting Error (m)		Northing Error (m)		R_T Error (m)	
	mean	RMS	mean	RMS	mean	RMS
Oct 3 Loc 1	-58	125	-87	133	-52	78
Oct. 24 Loc 1	-10	187	58	116	36	95
Oct. 3 Loc 2	185	237	-86	140	-111	165
Oct. 24 Loc 2	105	170	-195	222	-216	240
Oct. 24 Loc 3	65	274	-16	157	-24	185

The results presented indicate a horizontal position accuracy of 189 m to 287 m (67%). It was concluded in Chapter 7 that with HDOP $\leq$ 2 and in the absence of multipath, the Cellocate™ system is capable of 100 m (67%) positioning accuracy. The 89 m to 187 m of additional error here is mainly due to three factors. First is the accuracy of the site biases.

For these tests, the biases were measured on a separate date as opposed to being derived from the actual data from which position was estimated. The increased propagation distances are the second factor. In the static multipath test the propagation distances were on the order of 200 m and all LOS. Here the propagation distances vary from 322 m to 3 km and are often not LOS. This will obviously result in lower received SNR and therefore higher TOA estimation noise. The third factor is multipath. The lack of LOS propagation is one aspect and the presence of reflections is another. All three transmitter locations were next to busy roads and therefore subject to dynamic reflections from passing cars. Surrounding buildings also contribute static multipath.

The residual RMS values in Table 8.9 are a reflection of the propagation distance and, therefore, received SNR. For instance, the Renfrew cell site was most distant from locations 2 and 3. Correspondingly, the Renfrew residuals have a higher RMS than those of the other cell sites for those two locations. For location 1, Windwood is most distant and its residuals have the highest RMS. The Sheraton residual RMS values for Loc 2 and Loc 3 are roughly half of that for Loc 1. The propagation distance from Sheraton to Loc 1 is five times that to Loc 2 and three times that to Loc 2. On the other hand, the Franklin residual RMS values change very little even when the propagation distance is almost doubled. This may be due to greater sensitivity of the Franklin Cellocate™ receiver or that a LOS path always existed for the Franklin site.

Table 8.9   Statistics of Observation Residuals

	Franklin (m)		Renfrew (m)		Windwood (m)		Sheraton (m)	
	mean	RMS	mean	RMS	mean	RMS	mean	RMS
Oct 3 Loc 1	-1	79	-3	126	11	130	-6	78
Oct. 24 Loc 1	-35	79	76	173	-82	188	42	94
Oct. 3 Loc 2	1	70	-3	132	3	97	-1	34
Oct. 24 Loc 2	-10	84	26	160	-22	121	5	44

Oct. 24 Loc 3	-19	53	49	127	-52	121	22	49

The residuals for all data sets are highly correlated. According to Table 8.10, the magnitudes of the correlation coefficients range from 0.88 to 0.99. The average redundancy numbers for each cell site's observations are given in Table 8.11. Franklin and Sheraton have very low redundancy while Renfrew and Windwood have significantly higher redundancy. In any case, the redundancy numbers are such that the observations are still poorly controlled. When standardized residuals are calculated, they are found to be identical amongst the four cell sites as in Figure 7.16. Standardized residuals for the Oct. 3 Loc 1, Oct. 3 Loc 2, and Oct. 24 Loc 3 data sets are given in Figure 8.20. Note that a blunder is detected at epoch number 18 in the Oct. 3 Loc 1 data and this corresponds with a large positional error for that data epoch. A blunder is also detected at epoch 18 in the Oct. 24 Loc 3 data but does not correspond to a large positional error. Although identifying the epoch at which a blunder exists may be possible, it is impossible to determine which observation is at fault due to low redundancy.

Table 8.10   Residual Correlation Coefficients

	F / R	F / W	F / S	R / W	R / S	W / S
Oct. 3 Loc 1	-0.97	0.91	-0.94	-0.98	0.99	-0.99
Oct. 24 Loc 1	-0.99	0.98	-0.99	-0.99	0.99	-0.99
Oct. 3 Loc 2	-0.99	0.95	-0.96	-0.99	0.99	-0.99
Oct. 24 Loc 2	-0.99	0.97	-0.98	-0.99	0.99	-0.99
Oct. 24 Loc 3	-0.99	0.96	-0.88	-0.98	0.90	-0.97

Table 8.11   Average Observation Redundancy Numbers

	Franklin	Renfrew	Windwood	Sheraton
Oct. 3 Loc 1	0.07	0.32	0.46	0.15
Oct. 24 Loc 1	0.08	0.35	0.44	0.13
Oct. 3 Loc 2	0.14	0.52	0.30	0.04
Oct. 24 Loc 2	0.15	0.53	0.29	0.04

Oct. 24 Loc 3	0.07	0.42	0.43	0.08
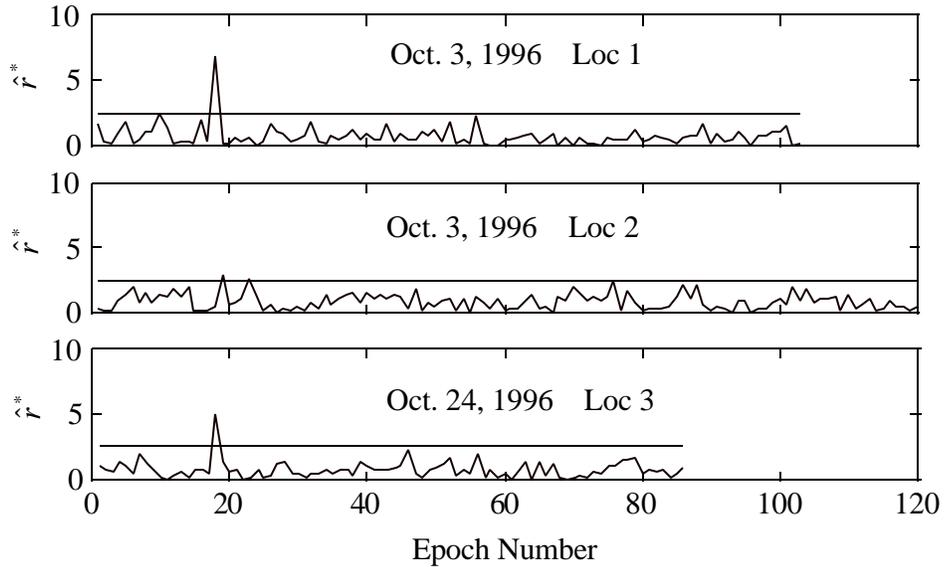
Figure 8.20   Standardized Residuals

8.4.3.2 Effect of MUSIC TOA Estimation

To illustrate the improvement in TOA estimation accuracy due to root MUSIC, position was estimated for all five data sets with TOA estimates obtained from the standard cross-correlation method described in section 3.6. The results may then be compared with those of root MUSIC. Other than the source of the TOA estimate, the position estimation process was identical to that in the previous section. The position performance measures are tabulated in Table 8.12. These may be compared with those of Table 8.7. The position estimates for the Oct. 3 Loc 1 data set are plotted in Figure 8.21 and should be compared to those of Figure 8.11.

Table 8.12   Position Results Using Full Correlation TOA Estimates

	DRMS (m)	Hor. Prec. (m)	67% (m)	95% (m)

Oct. 3 Loc 1	632	623	652	1124
Oct. 24 Loc 1	644	625	647	1129
Oct. 3 Loc 2	684	665	717	1013
Oct. 24 Loc 2	628	600	638	1053
Oct. 24 Loc 3	725	610	789	1078
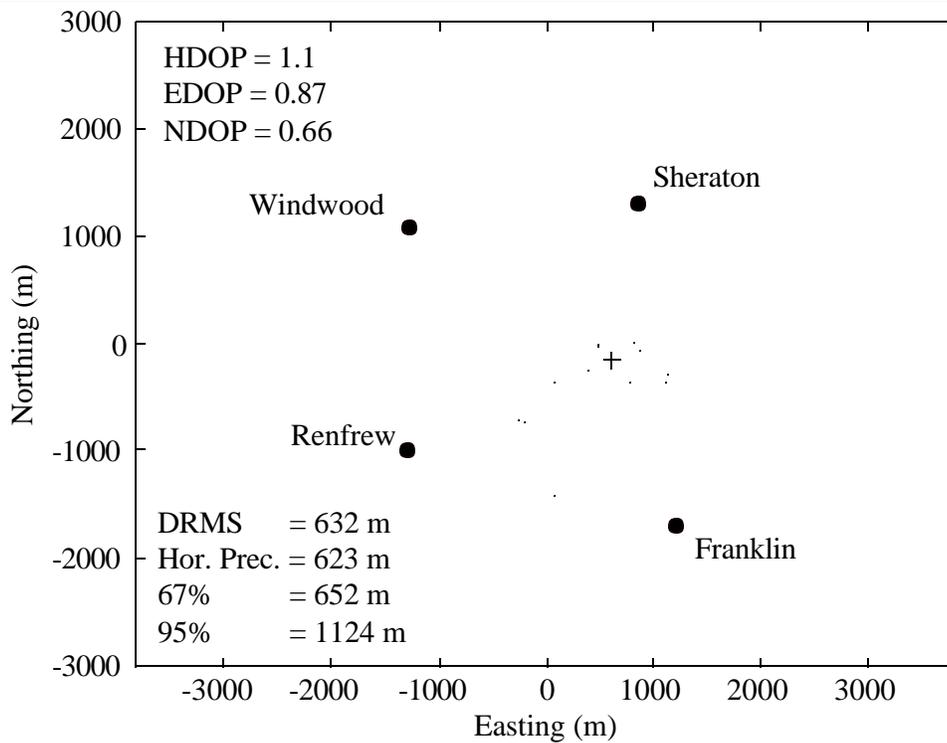
Figure 8.21   Position Results for Oct. 3 Loc 1 and Full Correlation TOA
                        Estimates

The 67th percentiles of Table 8.12 are anywhere from 2.2 to 3.5 times greater than those of Table 8.7 and 5 to 6 times greater than the FCC specification of 125 m. This clearly illustrates that traditional Fourier based correlation techniques provide insufficiently accurate TOA estimates for the purpose of positioning. Further processing by algorithms such as MUSIC is necessary to meet the FCC specification.

8.4.3.3 Effect of Averaging

The position results were averaged, as in Chapter 7, in order to determine the improvement in horizontal precision. Raw TOA observations, corrected by the measured average bias differences, were used. Figure 8.22 shows the results for the Oct. 3 Loc 1 data set while Figure 8.23 shows the same for the Oct. 3 Loc 2 data set. These figures are similar to Figure 7.21 in that they illustrate horizontal precision as a function of averaging time. The empirical results are again represented by the dotted lines whereas the theoretical results from (7.6) are represented by the solid lines.

The results of Figures 8.22 and 8.23 may be related to the results of Table 8.5. The integration time of the filter used to smooth the bias differences in Table 8.5 was 112 seconds. One would expect then that the standard deviation of the filtered bias differences could be related to the horizontal precision obtained after averaging for 112 seconds. For example, taking the average filtered standard deviation for the Oct. 3 Loc 1 data set in Table 8.5, dividing by $\sqrt{2}$, and multiplying by an HDOP of 1.1 results in a horizontal precision of 38 m. In Figure 8.22, the horizontal precision after an averaging time of 112 seconds is approximately 24 m. Although one might anticipate the two horizontal precisions to be equal, it is reasonable to expect a somewhat lower value from Figure 8.22 than that obtained from the results of Table 8.5. The standard deviations in Table 8.5 are calculated from the full data sets whereas the horizontal precisions in Figure 8.22 and 8.23 are calculated from subsets of the data sets (with the exception of the last one).
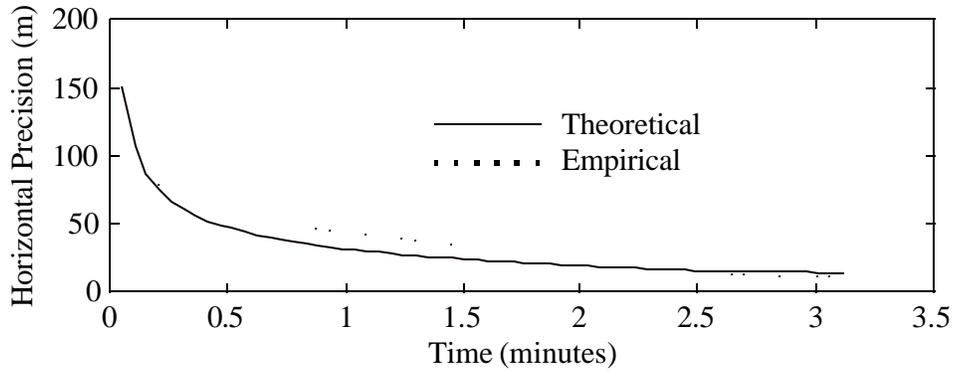
Figure 8.22   Effect of Averaging on Horizontal Precision

Oct. 3 Loc 1

The time interval between position estimates in both Figure 8.22 and Figure 8.23 is approximately 3 seconds. Thus, thirty seconds of averaging time corresponds to approximately 10 position fixes. Using 30 seconds of averaging the horizontal precision drops by a factor of 2.8 in Figure 8.22 and by a factor of 2.4 in Figure 8.23. For the other three data sets, the horizontal precision dropped by a factor of 1.9 to 2.4 after 30 seconds. After one minute, the drop was a factor anywhere from 2.8 to 3.6. If an infinite population is assumed, the finite population correction factor of (7.6) may be assumed to be unity. In that case, the theoretical number of fixes required to cut the horizontal precision in half is 4. To reduce the horizontal precision to one third, 9 fixes are required.

223
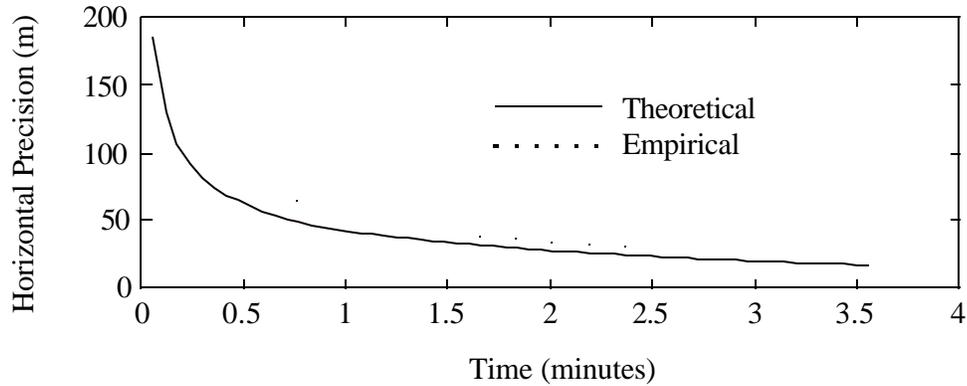
Figure 8.23   Effect of Averaging on Horizontal Precision

Oct. 3 Loc 2

The averaging time required to improve the horizontal precision will obviously depend on the instantaneous horizontal precision, which is a factor of TOA estimation error, SNR, and dynamic multipath, as well as the time between fixes. The Cellocate™ hardware is currently able to provide a TOA estimate every two seconds. To transmit the TOA estimates to a central site and combine them to estimate a position may require another 2 or 3 seconds. Therefore, if the network is able to page the cell phone every 2 seconds, and each cell phone precursor transmission results in a TOA estimate at four or more sell sites, 4 position fixes should be obtained in about 10 seconds. Averaging position estimates is, therefore, a very worthwhile exercise. This is important for the application of interest, enhanced 911 services. For a time interval of approximately 3 seconds between position fixes, averaging over 30 seconds will result in reducing the horizontal precision to approximately one third of its instantaneous value. Because time is of the essence in emergency situations, it is crucial that averaging times required for highly precise position, are short.

Note, however, that only precision improves with averaging. It is noise which is removed by averaging and any constant bias in the position estimates remains. Figures 8.22 and 8.23 show the effect of averaging on horizontal precision. Therefore, any bias in the positions will not be visible in Figures 8.22 and 8.23.

CHAPTER 9

CONCLUSIONS

9.1 TOA Estimation Simulations

The objective of the TOA estimation simulations was to determine the effect of multipath on TOA estimation by root MUSIC. Multipath was modeled as a 2-ray channel. During the course of the simulations it was found that a difference in the DC level of the correlation sequence, formed from the received data, from that of the reference correlation sequence, impacted TOA estimation accuracy. When the DC levels were matched, there was a 5% improvement in the TOA estimation accuracy.

The effect of multipath on TOA estimation by root MUSIC was investigated as a function of three multipath parameters: delay, phase, and amplitude. The mean root MUSIC TOA estimation error was found to be a function of the time delay between the LOS path and the second arrival. As the delay spread increased from zero, the mean TOA error also increased as expected. The mean error was not a function of SNR, but rather, was relatively constant over all noise levels. The levels of delay spread for suburban and urban environments were chosen to be 200 ns and 400 ns respectively. These values are consistent with empirical data. For a second arrival inphase with the LOS arrival but only of half amplitude, the mean TOA estimation error of root MUSIC was 20 m in the suburban case and 40 m in the urban case.

The standard deviation of the TOA estimation error was a function of the SNR - obviously decreasing as the SNR increased. However, standard deviation was virtually identical for all multipath delay spreads tested. At 15 dB, the standard deviation of the TOA estimation error was approximately 80 m and at 20 dB, approximately 45 m.

The phase of the second arrival, with respect to LOS, also affected TOA estimation. Both the mean and standard deviation of the TOA error were lowest for a multipath phase of 0° and highest for 180°. A phase difference of 180° between the two arrivals causes destructive interference and fading of the received signal amplitude. The effect on TOA estimation is, therefore, more pronounced at lower SNR values.

As expected, the smaller the amplitude of the second arrival with respect to the first, the smaller the effect on TOA estimation. For any value of multipath phase, the mean and standard deviation of the TOA estimation error increased as the amplitude of the second arrival increased. Again, this was particularly true at lower SNR values.

The simulated root MUSIC TOA estimation errors were found to be normally distributed.

9.2 Positioning Simulations

The purpose of the positioning simulations was to estimate the performance of Cellocate™ on a city-wide scale with an existing cellular cell site network. TOA estimation error models based on the results of Chapter 5 and multipath models based on empirical data were used to corrupt ranges from hypothetical cellular telephone locations to all cell sites. An SNR threshold was used to decide which cell sites would participate in the positioning process.

9.2.1 Geometry

The geometry of the Telus Mobility cellular network in Calgary, Alberta, was analyzed by calculating the HDOP for a dense grid of hypothetical cellular telephone locations. The area of best geometry (lowest HDOP) was the region just outside the 2 km radius urban zone of the city centre. The HDOP was found to be 1 or less in this area. Outside this first region to a radius of almost 12 km, was the next best area for geometry. In this region the HDOP was generally between 1 and 3. Within the 2 km city centre zone the HDOP was usually 5 or greater whereas at the edges of the 12 km radius the HDOP was generally greater than 8.

In general, the HDOP was lowest and geometry strongest in those areas where a large number of cell sites surrounded the point of interest. Along the edges of the 12 km radius, the number of cell sites is small and usually to one side of the point of interest. This begins to resemble the rural environment in which there are very few cell sites. For a TDOA system such as Cellocate™ to operate requires at least three cell sites to be involved. In a

rural environment three sites may not always be available. A possible solution is to install additional cell sites for positioning purposes only. These 'pseudo' cell sites need not be allocated voice channels for communication purposes. They could merely monitor the existing channels for positioning purposes.

On the whole, the geometry of the Telus Mobility network in Calgary is excellent. Of the points for which a solution was found, over 50% had an HDOP of 1 or less and 80% had an HDOP of 2 or less. Expressed as a percentage of the total number of grid points within a 12 km radius of the city centre, the values are 44% and 71% respectively. These results are based on the assumption that an average of 8 cell sites can 'hear' a cellular telephone. Although this average number is based on hand-off information from Telus Mobility, its accuracy is not certain. On the other hand, the simulations were performed for a total of 36 cell sites. Since the time of the simulations, the number of cell sites has increased and is ever increasing due to the need for more cellular capacity. As a result, the geometry results obtained here, if not already realistic, will become even more realistic, to the point of becoming conservative, as the density of cell sites increases.

9.2.2 Plane Intersection Simulation Results

Least squares requires an a priori position estimate to begin the iterative process. Plane intersection, a closed form algorithm which estimates position from TOA differences, was in general found to be capable of providing this position estimate. A simulation run was conducted over the entire 12 km radius test area using plane intersection to estimate horizontal position. The result was a horizontal error 67th percentile of 190 m. The DRMS was found to be 2784 m.

In general, the accuracy of plane intersection reported is sufficient for the purposes of providing an initial position estimate to LS. However, there may be cases when this is not true. Plane intersection requires at least 4 TOA observations for a horizontal position estimate whereas hyperbolic trilateration models require a very minimum of 3. When the number of TOA observations is only 3, some other source must be found for the initial position estimate. Secondly, due to the small distances involved in the cellular positioning

problem, it is possible for hyperbolic lines of position to cross each other in more than one place. These multiple intersections can occur very close to one another. Depending on the initial position estimate provided, it is possible that LS will converge to any of these multiple crossings. When these ambiguous solutions are close to one another, it is imperative that the initial position estimate be as close as possible to the true solution. For the two examples investigated, the accuracy of the initial position estimate was required to be 700 m in one case and 2 km in the other.

Fortunately, it is possible to detect the presence of two solutions with the bifurcation parameter. This parameter is easily computed for two hyperbola given the positions of the cell sites and the TOA differences. A positive value for the parameter indicates the presence of two solutions. Given the detection of two solutions, appropriate steps may be taken to minimize the possibility of converging to two separate solutions depending on the initial position used for LS. This may involve the use of a closed form algorithm of sufficient accuracy and not subject to solution bifurcation. Reducing the weight of, if not eliminating all together, the offending observations may also be required.

9.2.3 Least Squares Simulation Results

Using LS for positioning, the average 67th percentile of the instantaneous horizontal error was 119 m. This was based on 10 independent simulation runs. The average DRMS was found to be approximately 700 m because of a few outliers. If the assumptions and error models made in the simulations are correct, the Cellocate™ system is able to meet the FCC specification of 125 m (67%) for the Telus Mobility network in Calgary.

For each of the 10 simulation runs, there were a total of 7211 hypothetical cellular telephone locations within a 12 km radius of the city centre. On average, 10.2% of these points were visible from less than 4 cell sites. The positions of these points were not estimated due to the use of plane intersection for providing the initial position estimate. LS diverged for 1.65% of those points for which 4 or more TOA observations were available. The mean number of TOA observations was 8.2.

228

Within the 2 km radius urban zone, the 67th percentile was 327 m. In the 2 km to 12 km radius suburban zone, the 67th percentile was 116 m. The principal problem in the urban zone was geometry. Although at least 3 cell sites were visible from every hypothetical telephone location, the geometry was generally poor and this is reflected in the position results. In contrast, the geometry in the suburban zone was generally good with the exception of the city edges where a large number of hypothetical locations were visible from less than 4 cell sites. Therefore, the two zones suffer from different problems - poor geometry in the urban zone and a lack of coverage along the edges of the suburban zone.

The choice of the trilateration model was found to be essentially irrelevant. The 67th percentile was equal for all the models used. There were differences in a very few cases of poor geometry where numerical instabilities caused one model to converge and another to diverge or converge to a very different solution. The choice of model may then be a function of something other than solution accuracy. For instance, when residuals are required for the individual TOA observations, circular trilateration may be used without fear that accuracy is being compromised. For hyperbolic trilateration, sequential differencing is favored over reference differencing although the advantage may be very slight.

Running the same 10 simulation runs again, but without TOA estimation error, resulted in a 67th percentile of 52.3 m. This reflects the level of positioning error due to simulated multipath only. In this context, multipath refers to the probability that the first arrival does or does not follow a LOS path. The same simulation runs performed with TOA estimation error only, resulted in an accuracy of 99.1 m (67%). Any effort to improve positioning accuracy should concentrate on the improvement of TOA estimation since the multipath error added accounts for the lack of a LOS path and this cannot be changed. Positioning accuracy was found to improve until the standard deviation of the TOA estimation error was reduced to approximately 50 m. Any further improvement of the TOA estimation process was swamped out by the multipath effect of no LOS path.

As expected, geometry was found to significantly affect position estimation accuracy. As HDOP decreases, positioning accuracy improves. For those points with an HDOP of 2 or less, the horizontal position accuracy was 94 m (67%). This corresponds to

80% of the points for which a solution was computed. The conclusion is again made that, given the assumptions of the simulations, the geometry over the majority of the test area is sufficient to yield a position accuracy which meets the FCC specification. In an urban environment such as Calgary, Cellocate™ is capable of providing the position accuracy necessary for enhanced 911 purposes.

9.3 Static Multipath Field Test

9.3.1 TOA Estimation Noise

The static multipath field test was designed in order to estimate the TOA estimation noise in the Cellocate™ system. Removal of the long term drift from the TOA differences resulted in noise standard deviations ranging from 49 m to 113 m, with a mean of 79.5 m. In so doing, it is assumed that any static multipath is removed by the filtering and no dynamic multipath exists. The average standard deviation for an individual TOA estimate is, therefore, 79.5 m / $\sqrt{2}$ = 56.2 m. From this one may predict a horizontal precision of 56.2 m $\times$ HDOP.

The standard deviation of the noise varied from site to site. It could not be determined with the data collected in this particular test whether this was due to differences in the receivers themselves or differences in the physical sites. The propagation distances ranged from 190 m to 220 m suggesting that the received SNR should have been roughly equal for all receivers. However, differences in noise standard deviations were also found in the operational field tests. In both tests, one of the receivers displayed higher noise levels. One concludes that differences in terms of sensitivity exist between the four Cellocate™ receivers tested.

A long term drift was also present in the TOA differences. It is likely that this was due to temperature dependencies in the RF components. Filters with a group delay less dependent on temperature will rectify this problem.

9.3.2 Plane Intersection Position Results

Plane intersection was found to be very sensitive to the bias differences used to correct the TOA differences. Average bias differences derived from the data resulted in positioning accuracy and precision on the order of kilometres. This is obviously not adequate for the case at hand. If epoch by epoch lowpass filtered bias differences were used instead, a 67th percentile of 56 m was achieved. This is equivalent to the standard deviation of the TOA estimation noise and sufficient for an initial position for LS. This is true, however, only when there is no long term drift in the biases.

9.3.3 Least Squares Position Results

In terms of positioning accuracy, LS was less sensitive to the bias differences used. With average bias differences, LS achieved a horizontal error of 127 m (67%) and a horizontal precision of 143 m. Due to the use of average bias differences, the long term drift of the dominant bias difference was visible in each of the estimated parameters as well as each residual. This again emphasizes the need to prevent long term drift due to temperature dependence.

The observation residuals were found to be highly correlated with one another. This was due to low redundancy and a high degree of symmetry in the geometry. Calculation of each observation's redundancy number found the observations to be poorly controlled. As demonstrated in further simulations, it is impossible to detect a large error in one observation, when the observations are poorly controlled. When the number of observations was doubled, thereby changing the geometry as well as increasing the amount of redundancy, the observations became better controlled. Each observation's residuals then reflected the errors in that particular observation more accurately.

This was also found in the statistical testing of the residuals. For low redundancy, the statistical testing of each residual indicated the presence of a blunder. However, the blunder had only been introduced into one observation. It was, therefore, impossible to determine which observation contained the blunder. With additional redundancy, it was possible to identify the offending observation with statistical testing of the residuals. Therefore, residuals should only be used as indicators of individual observation quality when

the observations are well controlled. Redundancy numbers greater than 0.5 indicate relatively well controlled observations.

With the use of lowpass filtered bias differences, LS achieved a 67th percentile for horizontal position error of 105 m. The horizontal precision and DRMS were both 118 m. Recall that the expected horizontal precision for an HDOP of 1.95 was 110 m. The position results are unbiased, as indicated by the equivalence of DRMS and horizontal precision, and approximately Gaussian due to the agreement of the DRMS and 67th percentile. This was obviously expected since filtered bias differences, derived from the data, were used.

These results tend to confirm those of the simulations conducted in Chapter 6. In the absence of multipath and for good geometry (HDOP $\leq$ 2) Cellocate™ is capable of a positioning accuracy of approximately 100 m (67%). However, it must be stressed that this is dependent on the received SNR.

As expected, averaging of instantaneous position fixes results in an improvement in positioning performance. An instantaneous position fix was available every 8 seconds. Averaging 15 position fixes over 2 minutes resulted in a horizontal precision of 20 m. For an operational system, instantaneous positions will be available every 2 to 3 seconds. For the configuration of the static multipath test, averaging over periods of 30 seconds will give a horizontal precision of 20 m, a remarkable and critical result for emergency situations.

9.4 Operational Field Tests

9.4.1 CellocateÔ Receiver and Site Biases

The Cellocate™ receiver biases were measured in the laboratory by connecting each receiver, in turn, to a transmitter. Subtracting the TOT from the TOA gave a measurement of the receiver bias. The main conclusion from this test was, as previously mentioned, that the receiver biases are temperature dependent. Each receiver required 10 to 15 minutes after power-up before stabilizing to a constant bias.

When the Cellocate™ receivers were located in the Telus Mobility cell sites, the combined receiver and site bias was measured in a similar manner. With the transmitter located next to the cell site antenna, the TOT was again subtracted from the TOA. It is

assumed that the biases measured in this way contain no multipath effects. Before collecting the TOT and TOA data, the receivers were allowed to warm up. This resulted in no warm up drift and less long term drift compared to the biases of the static multipath test. The strictly controlled temperature in the cell sites, as well as allowing the receivers to warm up, helped to offset the temperature dependency of the RF filters in the receivers. New hardware designs include filters with group delay far less dependent on temperature.

9.4.2 Multipath Effects

From five static data sets covering three separate transmitter locations and two separate days, it was possible to estimate relative amounts of multipath. Differencing TOA estimates between any two cell sites and subtracting the known range difference, resulted in bias differences. These bias differences were then lowpass filtered to remove the TOA estimation noise. Subtracting the previously measured bias differences from these filtered bias differences gave an estimate of the multipath in the TOA differences.

From the above results it was possible to estimate the relative multipath effect between TOA observations for each of the five data sets. Assuming that the cell site judged to have the least multipath had in fact none, the absolute multipath effect on any TOA observation ranged from 30 m to 336 m. This agrees with the empirical data of Turin et al. (1972a) and the simulation multipath models of Chapter 6. One may therefore conclude that under the conditions encountered, the above method of estimating multipath is reasonable.

9.4.3 TOA Estimation Noise

An estimate of the TOA estimation noise was also made. This was done by subtracting the lowpass filtered bias differences from the raw bias differences. The level of the TOA estimation noise here was roughly 2.6 times higher than that of the static multipath test. This reflects the longer propagation paths here. The average TOA estimation noise on any individual TOA was 146 m.

9.4.4 Plane Intersection and Initial Position Accuracy

Plane intersection was again found to be very sensitive to the bias differences used to correct the TOA differences. Using lowpass filtered, epoch by epoch, bias differences, plane intersection achieved a position accuracy of 351 m to 982 m (67%). This level of accuracy resulted in the convergence of LS in 72% to 92% of all cases considered. However, accurate bias differences are required if plane intersection is to be used for providing initial positions.

9.4.5 Least Squares Position Results

Using the average measured bias differences, the position accuracy achieved with LS ranged from 189 m to 287 m (67%). This is 84 m to 182 m higher than that achieved in the static multipath test. One major reason for this difference is the propagation distances and hence, received SNR. As received SNR decreases, TOA estimation noise increases. The mean TOA estimation noise on any individual TOA was 56 m in the static multipath test. In the operational field tests this figure was 146 m.

A second major reason for the difference in positioning results between the static multipath test and the operational field tests is the bias differences used. The lowpass filtered, epoch by epoch bias differences used in the static multipath test accounted for the biases in the receivers, any long term drift in those biases, and any static multipath. The average measured biases used in the operational field test only account for the biases in the receivers and the cell site equipment. Any static multipath remains in the TOA observations. As indicated previously, these multipath effects ranged form 30 m to 336 m and manifested themselves in the biased position results.

Therefore, aside from maintaining good geometry, there are two major factors which determine whether Cellocate™ meets the FCC specification of 125 m (67%), namely received SNR and system calibration. Any effort to increase the received SNR will result in lower TOA estimation noise and consequently more accurate TOA estimates. Accurate system calibration will account for the biases of the receivers and the cell sites. If biases based on particular cellular telephone locations can be determined, their use will also remove much of the static multipath. A possibility for a working system is to estimate position in

two steps. In the first step, generic biases for the cell sites are used to estimate the telephone's position. Then, biases specific to the area in which the telephone is estimated to be in, may be used to generate a second more accurate position estimate.

As in the case of the static multipath test, averaging the position fixes had a significant impact on horizontal precision. Averaging some 10 position fixes was found to reduce the horizontal precision to about one third of its instantaneous value. In an operational system, 10 fixes should be available in 30 seconds or less. Averaging over approximately 30 seconds resulted in a horizontal precision of 50 m to 100 m for the five data sets considered. However, averaging does not reduce the effects of biases on position. From an emergency point of view, and for the purposes of dispatching emergency units, such a high level of precision may be critical.

9.5 Recommendations and Future Work

From this work recommendations may be made for an operational system as well as for future work. An important discovery made was the possibility of situations in which two solutions to the hyperbolic trilateration equations exist. These two solutions are often so close that either may be deemed reasonable. If the solutions are only 100 m apart, it may not be critical which is arrived at. However, remembering that the FCC requirement is 125 m (67%), it becomes very important to converge to the correct solution as the solutions become farther apart. As detailed in Chapter 6, it is possible to detect the presence of two solutions for a hyperbola pair. Therefore, it is recommended that an algorithm be developed to combine the observations is such a way as to minimize, if not eliminate, the possibility of two solutions in these cases.

Related to the problem of two possible solutions is the accuracy of the initial position provided to LS. Closed form algorithms such as plane intersection are the preferable source for the initial position. Future work should include an investigation of other closed form solutions in terms of accuracy and susceptibility to multiple solutions.

In all cases, position estimates should be averaged where possible to produce a more precise position estimate. Averaging reduces the effect of noise as well as any dynamic multipath due to motion of the telephone or reflecting objects.

In the event that relative biases instead of absolute biases are available to calibrate the system, circular trilateration with TOA differences is the recommended trilateration model. This will allow for statistical testing of individual TOA observations and consequently the detection of blunders. It would also allow for the calculation of observation redundancy numbers which indicate whether statistical testing of the residuals will be able to isolate the observation containing the blunder.

Important for an operational system is some sort of overall reliability parameter for the estimated position. This parameter should incorporate information from a number of sources such as the estimated parameter covariance matrix, statistical testing of the residuals, an internal and external reliability analysis, the bifurcation parameter, HDOP, SNR, and the amount of averaging performed. Emergency personnel who respond to a 911 call will want to know how much confidence they can place in the position estimate they are provided with.

Differences in height between the cell sites and the transmitting source were neglected in the field tests due to their small magnitude compared to the propagation distances involved. In an operational system, this will obviously not always be the case. In flat areas where it is probable that any telephone will be at or near a nominal height, the cell site antenna heights can be accounted for in the calibration of the cell site biases. In downtown areas, however, there can be large height differentials. The difference in height between a telephone at street level and a cell site located atop an office building can easily be on the order of 100 m and may account for the majority of the propagation distance. In such cases, the VDOP may be sufficiently low such that height could be estimated in addition to horizontal position. Cellocate™ should be tested in such environments in order to determine its ability to estimate the three dimensional position of a cellular telephone.

Cellocate™, as described herein, is configured for the AMPS cellular standard. As the analog world gives way to the digital, Cellocate™ must be adapted to work for other

standards. In North America, Code Division Multiple Access (CDMA) and Time Division Multiple Access (TDMA) (IS-136) are being implemented whereas in other parts of the world the Global System for Mobile Communications (GSM) is the dominant standard. Adapting Cellocate™ to these different standards will ensure that it remains a leading edge technology in the field of cellular telephone positioning.

REFERENCES

Associated Group, Inc., (1997). "Time Difference of Arrival Technology for Locating Narrowband Cellular Signals," on web page of TruePosition™, www.trueposition.com.

Associated Press, (1997). "Woman cried, prayed during 40-hour ordeal," Sioux Falls, South Dakota, Jan. 12.

Carter, G.C. (Editor) (1993). *Coherence and Time Delay Estimation*, IEEE Press, New York.

Chaffee, J.W., Abel, J.S. (1993). "Bifurcation of Pseudorange Equations," in *Proceedings of the 1993 National Technical Meeting*, San Francisco, California, January 20-22, 1993, The Institute of Navigation, pp. 203-211.

Dumont, L.R., (1994). "Super-Resolution of Discrete Arrivals in a Spread Spectrum System," M.Sc. Thesis, Department of Electrical and Computer Engineering, The University of Calgary, Calgary, Alberta, Canada.

FCC NEWSReport No. DC 96-52, June 12, 1996.

Fenton, P., Falkenberg, B., Ford, T., Ng, K., Van Dierendonck, A.J., (1991). "Novatel's GPS Receiver, The High Performance OEM Sensor of the Future," in *Proceedings of ION GPS-91,* Albuquerque, New Mexico, September 9-13, pp. 49-58.

Goud, P.A., (1991). "A Spread Spectrum Radiolocation Technique and its Application to Cellular Radio," M.Sc. Thesis, Department of Electrical and Computer Engineering, The University of Calgary, Calgary, Alberta, Canada.

Hata, M., (1980). "Empirical Formula for Propagation Loss in Land Mobile Radio Services," *IEEE Transactions on Vehicular Technology*, Vol. VT-29, No. 3, pp. 317-325, August.

Haykin, S. (1989). *An Introduction To Analog and Digital* Communications, John Wiley & Sons, Inc., New York.

Haykin, S. (1991). *Adaptive Filter Theory, Second Edition,* Prentice Hall, Englewood Cliffs, NJ.

Hofmann-Wellenhof, B., Lichtenegger, H., Collins, J., (1993). *GPS Theory and Practice, Second Edition,* Springer-Verlag, Wien New York.

Hurst, G.C., (1989). "A New AVL System Developed in Australia," in *Proceedings of First Vehicle Navigation and Information Systems Conference*, pp. A60-A62.

Kay, S.M., (1993). *Fundamentals of Statistical Signal Processing*, PTR Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Kennedy, J.P., Roller, C.D., Hooper, R.W., (1994). "Passive High Accuracy Geolocation System and Method," United States Patent Number 5,317,323, May 31.

Kennedy, J., Sullivan, M.C., (1995). "Direction Finding and 'Smart Antennas' using Software Radio Architectures," *IEEE Communications Magazine,* May.

Klukas, R., (1993). "Angle of Arrival Estimation in the Outdoor Radio Environment," M.Sc. Thesis, Department of Electrical and Computer Engineering, The University of Calgary, Calgary, Alberta, Canada.

Klukas, R., Lachapelle, G., Fattouche, G., Borsodi, A., Astridge, M., (1996). "A System to Position Cellular Telephones Using GPS Time Synchronization," in *Proceedings of Third World Congress on Intelligent Transport Systems,* Orlando, Florida, Oct. 14-18.

Krakiwsky, E.J., (1990). The Method of Least Squares: A Synthesis of Advances, UCGE Report Number 10003, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, Canada.

Lachapelle, G., Cannon, M.E., Lu, G., (1992). "High-Precision GPS Navigation with Emphasis on Carrier-Phase Ambiguity Resolution," *Marine Geodesy*, Vol. 15, pp. 253-269.

Lachapelle, G., (1993). Navstar GPS Theory and Applications, ENGO 625 Lecture Notes, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, Canada.

Lee, W.C.Y., (1982). *Mobile Communications Engineering,* McGraw-Hill Book Company, New York.

Leick, A., (1995). *GPS Satellite Surveying, Second Edition,* John Wiley & Sons, Inc., New York.

Mackenzie, A.P., (1985). "Design and Assessment of Horizontal Survey Networks," M.Sc. Thesis, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, Canada.

Matlab[®], (1994a). *Statistics Toolbox User's Guide*, The MathWorks Inc., Natick, Massachusetts, pp. 1-6.

Matlab®, (1994b). *Reference Guide*, The MathWorks Inc., Natick, Massachusetts, pp. 27-32.

Maybeck, P.S., (1994). *Stochastic Models, Estimation, and Control, Volume 1,* Navtech Book & Software Store.

Microwaves & RF (1995). News Column in *Microwaves & RF,* pg. 46, March.

Mikhail, E.M., Ackermann, F., (1976). *Observations and Least Squares*, IEP - A Dun-Donnelley Publisher, New York.

Morley, G.D., (1995). "Improved Location Estimation for Automatic Vehicle Monitoring with Multipath-Corrupted Range Measurements," M.Eng. Project Report, Department of Electrical Engineering, The University of Alberta, Edmonton, Alberta, Canada.

Morley, T.G., (1997). "Augmentation of GPS with Pseudolites in a Marine Environment," M.Sc. Thesis, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, Canada.

National Research Council (1995). "The Global Positioning System, A Shared National Asset, Recommendations for Technical Improvements and Enhancements," National Academy Press, Washington, D.C.

Neter, P., Wasserman, W., Whitmore, G.A., (1972). *Applied Statistics*, Allyn and Bacon, Inc., Boston.

NovAtel Communications Ltd., (1994). *GPSCardÔ Command Descriptions Manual*, Publication Number OM-20000008, Revision Level 1.0, Software Version 2.2/3.2, NovAtel Communications Ltd., Calgary, Alberta, Canada.

Pillai, S.U., Kwon, B.H., (1989). "Forward/Backward Spatial Smoothing Techniques for Coherent Signal Identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 37, No. 1, pp. 8-14, January.

Proakis, J.G. (1989). *Digital Communications*, Second Edition, McGraw-Hill Book Company, New York.

Rao, B.D., Hari, K.V.S., (1989). "Performance Analysis of root MUSIC," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 37, No. 12, pp. 1939-1949, 1989.

Roberts, D.W., (1993). "Accuracy, Reliability and Repeatability of the 1991 Caucasia Global Positioning System Experiment," M.Sc. Thesis, Department of Geomatics Engineering, The University of Calgary, Calgary, Alberta, Canada.

Roth, P.R. (1971). "Effective Measurements using Digital Signal Analysis," *IEEE Spectrum*, Vol. 8, pp. 62-70, April.

Sakagami, S., Aoyama, S., Kuboi, K., Shirota, S., Akeyama, A., (1992). "Vehicle Position Estimates by Multibeam Antennas in Multipath Environments," *IEEE Transactions on Vehicular Technology*, Vol. 41, No. 1, pp. 63-68, February.

Sanders, A Lockheed Martin Company, Private Communication, March 25, 1996.

Schmidt, R.O. (1972). "A New Approach to Geometry of Range Difference Location," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 8, No. 6, pp. 821-835, November.

Schmidt, R.O. (1986). "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antennas and Propagation*, Vol. AP-34, No. 3, pp. 276-280, March.

Schmidt, R.O. (1996). "Least Squares Range Difference Location," *IEEE Transactions on Aerospace and Electronic Systems,* Vol. 32, No. 1, pp. 234-241, January.

Shan, T., Wax, M., Kailath, T., (1985). "On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 33, No. 4, pp. 806-811, August.

Smith, J.O., Abel, J.S., (1987). "Closed-Form Least-Squares Source Location Estimation from Range-Difference Measurements," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. 35, No. 12, pp. 1661-1669, December.

Spilker, Jr., J.J., (1980). "GPS Signal Structure and Performance Characteristics," in *Global Positioning System, Volume I,* The Institute of Navigation, Washington, D.C., pp. 29-54.

Steeves, R.R., Fraser, C.S., (1987). "Statistical Post-Analysis of Least Squares Adjustment Results," *Papers for the CISM Adjustment and Analysis Seminars*, Second Edition, Canadian Institute of Geomatics, Krakiwsky, E.J., Editor, pp. 182-210.

Stilp, L.A., Knight, C.A., Webber, J.C., (1994). "Cellular Telephone Location System," United States Patent Number 5,327,144, July 5.

Swokowski, E.W., (1983). *Calculus with Analytic Geometry*, Alternate Edition, Prindle, Weber & Schmidt, Boston, Massachusetts.

Turin, G.L., Clapp, F.D., Johnston, T.L., Fine, S.B., Lavry, D., (1972a). "A Statistical Model of Urban Multipath Propagation," *IEEE Transactions on Vehicular Technology*, Vol. VT-21, No. 1, pp. 1-9, February.

Turin, G.L., Jewell, W.S., Johnston, T.L., (1972b). "Simulation of Urban Vehicle-Monitoring Systems," *IEEE Transactions on Vehicular Technology*, Vol. VT-21, No. 1, pp. 9-16, February.

Van Dierendonck, A.J., Fenton, P., Ford, T., (1992). "Theory and Performance of Narrow Correlator Spacing in a GPS Receiver," *Navigation: Journal of the Institute of Navigation,* Vol. 39, No. 3, pp. 265-283.

Van Dierendonck, A.J., Melton, W.C., (1983). "Applications of Time Transfer Using NAVSTAR GPS," in *Global Positioning System, Volume II,* The Institute of Navigation, Washington, D.C., pp. 133-146.

Van Nee, D.J.R., (1991). "Multipath Effects on GPS Code Phase Measurements," in *Proceedings of ION GPS-91,* Albuquerque, New Mexico, September 11-13, pp. 915-924.

Van Nee, D.J.R., (1992). "GPS Multipath and Satellite Interference," in *Proceedings of ION 48th Annual Meeting,* Washington, DC, June 29 - July 1, pp. 167-177.

Van Valkenburg, M.E., (1964). *Network Analysis,* Second Edition, Prentice-Hall, Inc., Englewood Cliffs, N.J.

Warren, W.T., Whitten, J.R., Anderson, R.E., Merigo, M.A., (1972). "Vehicle Location System Experiment," *IEEE Transactions on Vehicular Technology*, Vol. VT-21, No. 3, pp. 92-101, August.

Yamada, H., Ohmiya, M., Ogawa, Y., Itoh, K., (1991). "Superresolution Techniques for Time-Domain Measurements with a Network Analyzer," *IEEE Transactions on Antennas and Propagation*, Vol. 39, No. 2, pp. 177-183, February.