UNIT 3

COORDINATE SYSTEMS AND SOURCES OF ERROR

SYLLABUS

Coordinate Systems
Geodetic reference systems
Earth-Centered Inertial Coordinate System
Earth-Centered Earth-Fixed Coordinate System
World Geodetic System
Indian Geodetic System (IGS)

Sources of error in GNSS
Satellite and Receiver clock errors
Ephemeris error
Multipath error
Atmospheric errors
Hardware bias error
Pseudorange error budget
Effects of Satellite Outages on GPS Availability

THE EARTH CENTRED INERTIAL (ECI) FRAME

Earth Centred Inertial frame has its origin right at the centre of the earth, however it is not fixed to the earth.

Although this frame has its origin at the centre of the earth, but it does not rotate with the earth.

The fundamental plane contains the equator and the positive X-axis points in the vernal equinox direction.

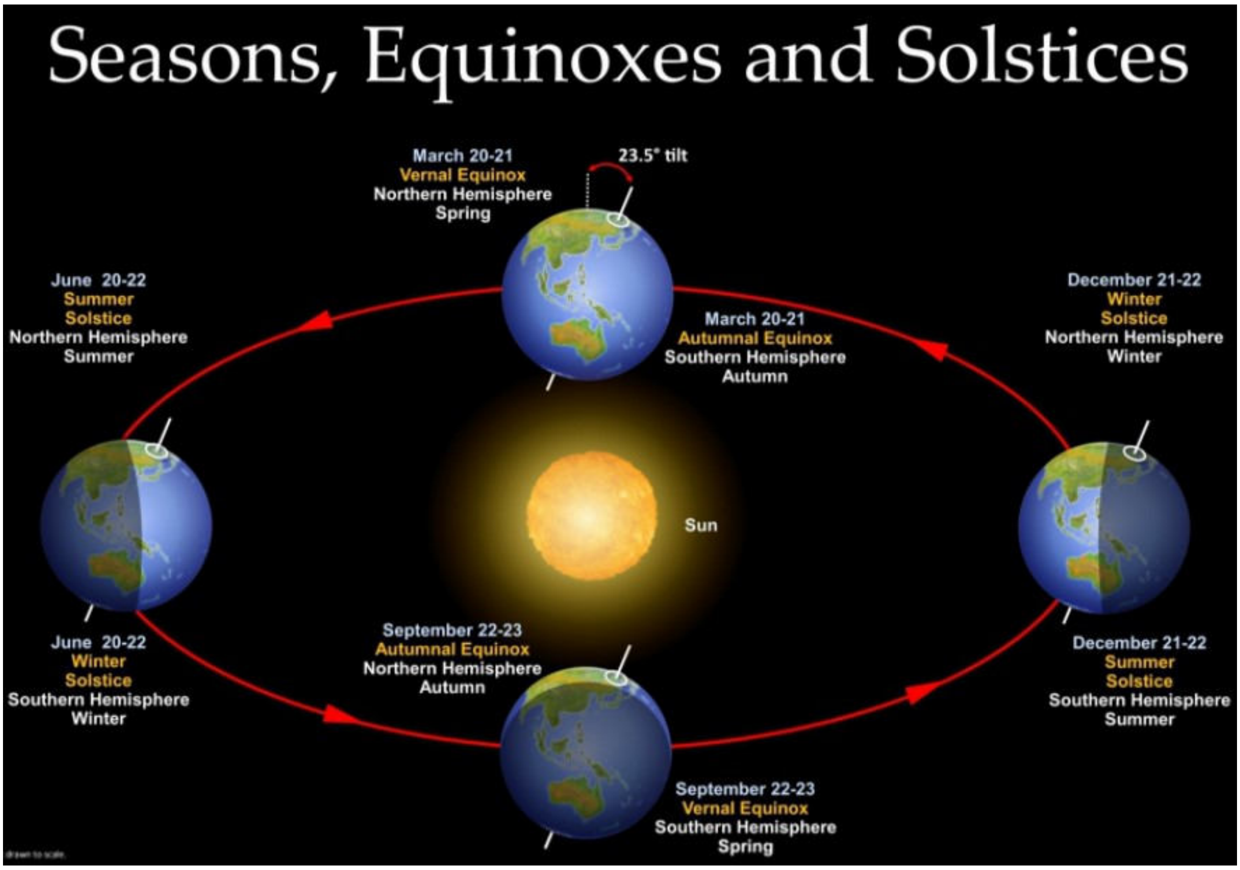

Fig. 3.1 Equinoxes and Solstices

The Z-axis points in the direction of the geographical North Pole and the Y axis consequently completes the right hand set of co-ordinate axes.

For objects in space, the equations of motion that describe orbital motion are simpler in a non-rotating frame such as ECI.

The ECI frame is also useful for specifying the direction toward celestial objects.

ECI coordinate frames are not truly inertial since the Earth itself is accelerating as it travels in its orbit about the Sun.

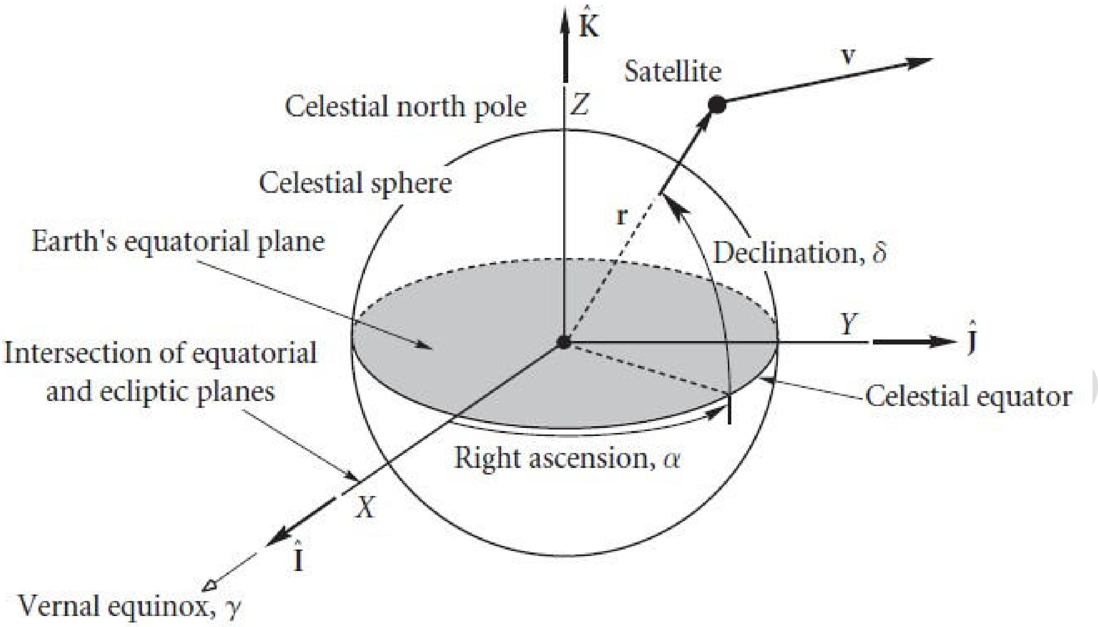

Fig. 3.2 ECI Frame

EARTH-CENTERED, EARTH-FIXED (ECEF)

The Earth-Centered, Earth-Fixed (ECEF) coordinate system is also known as **the "conventional terrestrial" coordinate system**.

It is a simple Cartesian coordinate system with the center of the earth at it's origin.

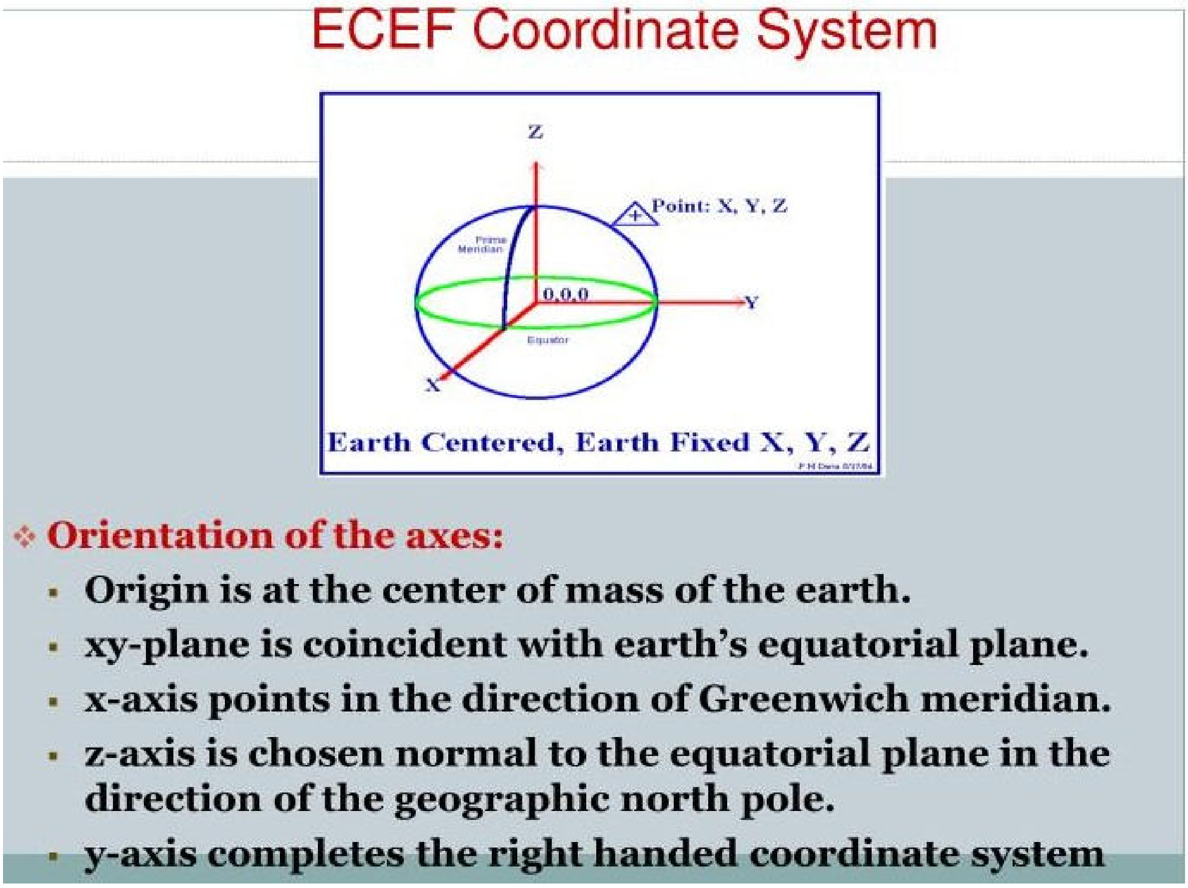

Fig. 3.3 ECEF Coordinate System

To represent the positions and velocities of terrestrial objects, it is convenient to use ECEF coordinates or latitude, longitude, and altitude.

ECEF and ECI coordinate frames have their origins at the Earth's Center of Mass.

ECI is called "Inertial" where as the Earth Centered, Earth Fixed (ECEF) frame rotates wrt inertial space to remain fixed to the surface of the Earth.

WORLD GEODETIC SYSTEM 84 (WGS 84)

The World Geodetic System 1984 (WGS84) is a datum featuring coordinates that change with time

WGS84 is defined and maintained by the United States National Geospatial-Intelligence Agency (NGA).

It is a global datum, which means that coordinates change over time for objects which are fixed in the ground.

The continuous ground movement means that even in the absence of earthquakes and other localised land movements, WGS84 coordinates are constantly changing.

These are often referred to as dynamic or kinematic coordinates.

Therefore it is important that coordinates in terms of WGS84 have a time associated with them, especially where the best levels of accuracy are required.

WGS84 is an accurate system as its center is estimated to be only + or – 2 meters away from the center of gravity of the Earth.

The system is therefore most suitable for higher defense and scientific applications.

INDIAN GEODETIC SYSTEM

India and other countries of the world made measurements in their countries and defined reference surface to serve as Datum for mapping.
In India the reference surface was defined by Sir George Everest, who was Surveyor General of India from 1830 to 1843.
It has served as reference for all mapping in India.
Indian system can be called Indian Geodetic System as all coordinates are referred to it.
The reference surface was called Everest Spheroid.

The initial point for mapping on the surface of the Earth was chosen at Kalyanpur in Central India.
Center of Everest Spheroid is about a km away from the center of gravity of the Earth; hence it is non-geocentric.
Indian Geodetic Datum is based on Everest Spheroid as Reference Surface and Kalyanpur in Central India as i
Center of this reference surface is estimated to be about 1 km away from the center of gravity of the Earth.
The datum is thus a local datum and in error.
Scientific and Defense studies of vital National importance cannot be based on such a system.
It is therefore extremely necessary that the Indian Geodetic Datum should be redefined at the earliest.

By Dr.Swapna Raghunath
Professor, Dept. of ECE

Vertical Datum for Heights in India was chosen as the Mean Sea Level at a group of nine tidal observatories situated at Indian ports.

Level network in India is of moderate to high precision at different places.

ONBOARD CLOCK ERRORS

Timing of the signal transmission from each satellite is directly controlled by its own atomic clock without any corrections applied.

This time frame is called *space vehicle* (SV) *time*.

Although the atomic clocks in the satellites are highly accurate, errors can be large enough to require correction.

Correction is needed partly because it would be difficult to directly synchronize the clocks closely in all the satellites.

Instead, the clocks are allowed some degree of relative drift that is estimated by ground station observations and is used to generate clock correction data in the GPS navigation message.

When SV time is corrected using this data, the result is called *GPS time*.

The time of transmission used in calculating pseudoranges must be in GPS time, which is common to all satellites.

The onboard clock error is typically less than 1 ms and varies slowly.

This permits the correction to be specified by a quadratic polynomial in time whose coefficients are transmitted in the navigation message.

The correction has the form

After the correction has been applied, the residual error in GPS time is typically less than a few nano seconds, or about 1 m in range.

RECEIVER CLOCK ERRORS

Because the navigation solution includes a solution for receiver clock error, the requirements for accuracy of receiver clocks is far less stringent than for the GPS satellite clocks.

In fact, for receiver clocks short-term stability over the pseudorange measurement period is usually more important than absolute frequency accuracy.

In almost all cases such clocks are quartz crystal oscillators with absolute accuracies in the 1–10 ppm range over typical operating temperature ranges.

When properly designed, such oscillators typically have stabilities of 0.01–0.05 ppm over a period of a few seconds.

Receivers that incorporate receiver clock error in the Kalman filter state vector need a suitable mathematical model of the crystal clock error.

In this model the clock error consists of a bias (frequency) component and a drift (time) component.

The frequency error component is modeled as a random walk produced by integrated white noise.

The time error component is modeled as the integral of the frequency error after additional white noise (statistically independent from that causing the frequency error) has been added to the latter.

In the model the key parameters that need to be specified are the power spectral densities of the two noise sources, which depend on characteristics of the specific crystal oscillator used.

EPHEMERIS DATA ERRORS

Small errors in the ephemeris data transmitted by each satellite cause corresponding errors in the computed position of the satellite (here we exclude the ephemeris error component of SA, which is regarded as a separate error source).

Satellite ephemerides are determined by the master control station of the GPS ground segment based on monitoring of individual signals by four monitoring stations.

Because the locations of these stations are known precisely, an “inverted” positioning process can calculate the orbital parameters of the satellites as if they were users.

This process is aided by precision clocks at the monitoring stations and by tracking over long periods of time with optimal filter processing.

Based on the orbital parameter estimates thus obtained, the master control station uploads the ephemeris data to each satellite, which then transmits the data to users via the navigation data message.

Errors in satellite position when calculated from the ephemeris data typically result in range errors less than 1 m.

Improvements in satellite tracking will undoubtedly reduce this error further.

THE MULTIPATH ERROR

Multipath propagation of the GPS signal is a dominant source of error in differential positioning.

Objects in the vicinity of a receiver antenna (notably the ground) can easily reflect GPS signals, resulting in one or more secondary propagation paths.

These secondary-path signals, which are superimposed on the desired direct-path signal, always have a longer propagation time and can significantly distort the amplitude and phase of the direct-path signal.

Errors due to multipath cannot be reduced by the use of differential GPS, since they depend on local reflection geometry near each receiver antenna.

In a receiver without multipath protection, C/A-code ranging errors of 10 m or more can be experienced.

Multipath can not only cause large code ranging errors but also severely degrade the ambiguity resolution process required for carrier phase ranging such as that used in precision surveying applications.

Multipath propagation can be divided into two classes: static and dynamic.

For a stationary receiver, the propagation geometry changes slowly as the satellites move across the sky, making the multipath parameters essentially constant for perhaps several minutes.

However, in mobile applications there can be rapid fluctuations in fractions of a second.
Therefore, different multipath mitigation techniques are generally employed for these two types of multipath environments.

In a receiver not designed expressly to handle multipath, the resulting cross-correlation function will now have two superimposed components, one from the direct path and one from the secondary path.

The result is a function with a distortion depending on the relative amplitude, delay, and phase of the secondary-path signal (in-phase secondary path or an out-of-phase secondary path).

The location of the peak of the function is displaced from its correct position, resulting in a pseudorange error.

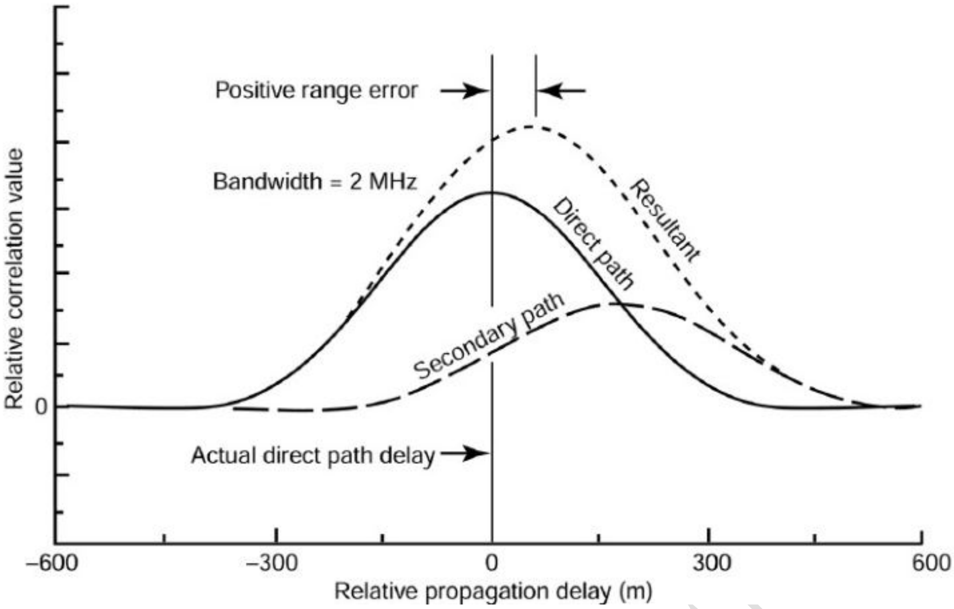

Figure 3.4 Multipath Error with Positive Error Range

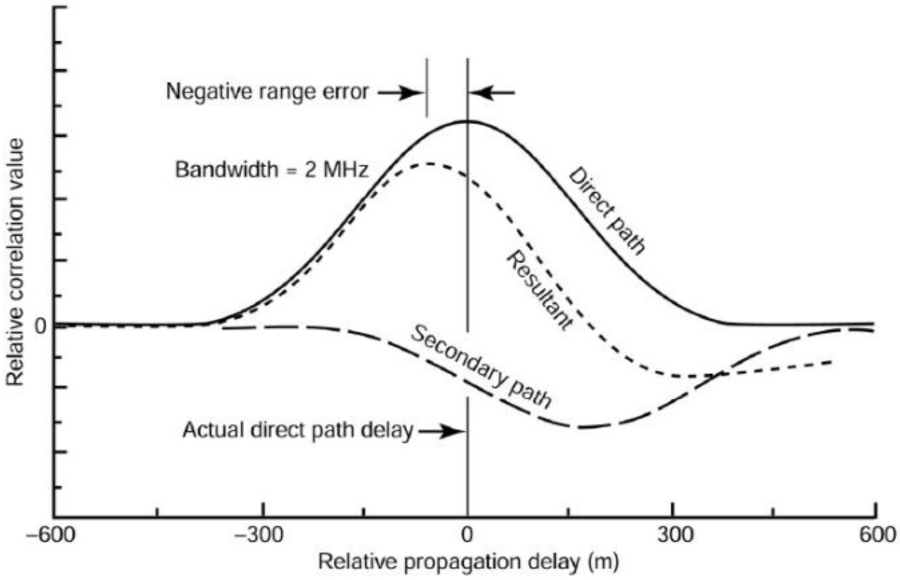

Figure 3.5 Multipath Error with Negative Error Range

In earlier receivers the magnitude of pseudorange error caused by multipath can be 70–80 m.

ATMOSPHERIC ERRORS

IONOSPHERIC PROPAGATION ERRORS

The ionosphere, which extends from approximately 50 to 1000 km above the surface of the earth, consists of gases that have been ionized by solar radiation.

The ionization produces clouds of free electrons that act as a dispersive medium for GPS signals in which propagation velocity is a function of frequency.

A particular location within the ionosphere is alternately illuminated by the sun and shadowed from the sun by the earth in a daily cycle; consequently the characteristics of the ionosphere exhibit a diurnal variation in which the ionization is usually maximum late in mid afternoon and minimum a few hours after midnight.

Additional variations result from changes in solar activity.

The primary effect of the ionosphere on GPS signals is to change the signal propagation speed as compared to that of free space.

A curious fact is that the signal modulation (the code and data stream) is delayed, while the carrier phase is advanced by the same amount.

Thus the measured pseudorange using the code is larger than the correct value, while that using the carrier phase is equally smaller.

The magnitude of either error is directly proportional to the total electron content (TEC) in a tube of 1 m2 cross section along the propagation path.

The TEC varies spatially, due to spatial nonhomogeneity of the ionosphere.

Temporal variations are caused not only by ionospheric dynamics but also by rapid changes in the propagation path due to satellite motion.

The path delay for a satellite at zenith typically varies from about 1 m at night to 5–15 m during late afternoon.

At low elevation angles the propagation path through the ionosphere is much longer, so the corresponding delays can increase to several meters at night and as much as 50 m during the day.

Since ionospheric error is usually greater at low elevation angles, the impact of these errors could be reduced by not using measurements from satellites below a certain elevation mask angle.

However, in difficult signal environments, including blockage of some satellites by obstacles, the user may be forced to use low elevation satellites.

Mask angles of 5◦–7.5◦ offer a good compromise between the loss of measurements and the likelihood of large ionospheric errors.
The L1-only receivers in non differential operation can reduce ionospheric pseudorange error by using a model of the ionosphere broadcast by the satellites, which reduces the uncompensated ionospheric delay by about 50% on the average.

During the day errors as large as 10 m at mid latitudes can still exist after compensation with this model and can be much worse with increased solar activity. Other recently developed models offer somewhat better performance.

However, they still do not handle adequately the daily variability of the TEC, which can depart from the modeled value by 25% or more.

The L1/L2 receivers in non differential operation can take advantage of the dependence of delay on frequency to remove most of the ionospheric error.

A relatively simple analysis shows that the group delay varies inversely as the square of the carrier frequency.

This can be seen from the following model of the code pseudorange measurements at the L1 and L2 frequencies:

$$\rho_i = \rho \pm \frac{k}{f_i^2},$$

where ρ is the error-free pseudorange,
ρi is the measured pseudorange,
k is a constant that depends on the TEC along the propagation path.

The subscript $i = 1, 2$ identifies the measurement at the L1 or L2 frequencies, respectively,

and the plus or minus sign is identified with respective code and carrier phase pseudorange measurements.

The two equations can be solved for both ρ and k.

$$\rho = \frac{f_1^2}{f_1^2 - f_2^2}\rho_1 - \frac{f_2^2}{f_1^2 - f_2^2}\rho_2,$$

The solution for ρ for code pseudorange measurements is where $f1$ and $f2$ are the L1 and L2 carrier frequencies, respectively,
and $\rho1$ and $\rho2$ are the corresponding pseudorange measurements.

With differential operation ionospheric errors can be nearly eliminated in many applications, because ionospheric errors tend to be highly correlated when the base and roving stations are in sufficiently close proximity.

With two L1-only receivers separated by 25 km, the unmodeled differential ionospheric error is typically at the 10–20-cm level.

TROPOSPHERIC PROPAGATION ERRORS

At 100 km separation this can increase to as much as a meter.

Additional error reduction using an ionospheric model can further reduce these errors by 25–50%.

The lower part of the earth's atmosphere is composed of dry gases and water vapor, which lengthen the propagation path due to refraction.

The magnitude of the resulting signal delay depends on the refractive index of the air along the propagation path and typically varies from about 2.5 m in the zenith direction to 10–15 m at low satellite elevation angles.

The troposphere is non dispersive at the GPS frequencies, so that delay is not frequency dependent.

In contrast to the ionosphere, tropospheric path delay is consequently the same for code and carrier signal components.

Therefore, this delay cannot be measured by utilizing both L1 and L2 pseudorange measurements, and either models and/or differential positioning must be used to reduce the error.

The refractive index of the troposphere consists of that due to the dry-gas component and the water vapor component, which respectively contribute about 90% and 10% of the total.

Knowledge of the temperature, pressure, and humidity along the propagation path can determine the refractivity profile, but such measurements are seldom available to the user.

However, using standard atmospheric models for dry delay permits determination of the zenith delay to within about 0.5 m and with an error at other elevation angles that approximately equals the zenith error times the cosecant of the elevation angle.

These standard atmospheric models are based on the laws of ideal gases and assume spherical layers of constant refractivity with no temporal variation and an effective atmospheric height of about 40 km.

Estimation of dry delay can be improved considerably if surface pressure and temperature measurements are available, bringing the residual error down to within 2–5% of the total.

The component of tropospheric delay due to water vapor (at altitudes up to about 12 km) is much more difficult to model, because there is considerable spatial and temporal variation of water vapor in the atmosphere.

Fortunately, the wet delay is only about 10% of the total, with values of 5–30 cm in continental midlatitudes.

Despite its variability, an exponential vertical profile model can reduce it to within about 2–5 cm.

In practice, a model of the standard atmosphere at the antenna location would be used to estimate the combined zenith delay due to both wet and dry components.
Such models use inputs such as the day of the year and the latitude and altitude of the user.

The delay is modeled as the zenith delay multiplied by a factor that is a function of the satellite elevation angle.

At zenith, this factor is unity, and it increases with decreasing elevation angle as the length of the propagation path through the troposphere increases.

Typical values of the multiplication factor are 2 at 30◦ elevation angle, 4 at 15◦, 6 at 10◦, and 10 at 5◦.

The accuracy of the model decreases at low elevation angles, with decimeter level errors at zenith and about 1 m at 10◦ elevation.

Although a GPS receiver cannot measure pseudorange error due to the troposphere, differential operation can usually reduce the error to small values by taking advantage of the high spatial correlation of tropospheric errors at two points within 100–200 km on the earth's surface.

However, exceptions often occur when storm fronts pass between the receivers, causing large gradients in temperature, pressure, and humidity.

HARDWARE BIAS ERROR

GNSS hardware biases appear in code and phase observations, and originates both from the receiver and satellite hardware. The presence of biases in GNSS observations might affect the accuracy in precise GNSS positioning applications, and might also be of relevance in other GNSS applications. They may also be a cause of incompatibility between different receiver types or GNSS constellations.

GNSS ERROR BUDGET

The understanding and management of errors is indispensable for finding the true geometric range (ρ) between a satellite and a receiver from a pseudorange (p).

$p = \rho + d\rho + c(dt - dT) + d_{ion} + d_{trop} + \varepsilon_{mp} + \varepsilon_p$ (pseudorange)

This equation includes environmental and physical limitations called *range biases*.

p = pseudorange

ρ = true geometric range

$d\rho$ = orbital errors

$(dt-dT)$ = clock errors

dion = ionospheric error

dtrop = tropospheric error

εmp = multipath

εp = receiver noise

Atmospheric errors are among the biases; two are the ionospheric effect, d_{ion}, and the tropospheric effect, d_{trop}.

Other biases, clock errors symbolized by *(dt-dT)* and receiver noise, ε_{ρ}, multipath, $\varepsilon_{m\rho}$, and orbital errors, d_{ρ}, are unique to satellite surveying methods.

Each of these biases comes from a different source.

They are each independent of one another but they combine to obscure the true geometric range.

There are many more elements, errors, or biases that contaminate the pseudorange— the satellite orbital errors, the ephemeris errors, etc.

Nevertheless, all are part of the error budget.

Table 3.1 GNSS Pseudorange Error Budget

GNSS Pseudorange Error Budget

ERROR SOURCE	ERROR CONTRIBUTION (m RMS)
Orbital	2.5
Satellite Clock	2
Receiver Noise	0.3
Ionospheric	5
Tropospheric	0.5
Multipath	1
Total	*11.3*

EFFECTS OF SATELLITE OUTAGES ON GPS AVAILABILITY

Being a part of the national infrastructure, satellite navigation systems considerably improve the operation of the existing and allow for the development and deployment of entirely new technical, economic and social systems.

However, the limitations and vulnerabilities of satellite navigation systems add their contribution to the risk budget of all systems that utilise satellite navigation services.

The American Global Positioning System (GPS) and Russian Glonass are currently the only fully operational satellite navigation systems, with Beidou and Galileo gradually building their infrastructures.

Satellite navigation systems are carefully planned and, with their eventual integration into the Global Navigation Satellite System (GNSS), will provide a reliable system capable to offer good satellite signal availability and satellite visibility.

Still, the satellite outages remain a potential cause of risk concerns.

Satellite navigation systems, like any other technical system, suffer from intrinsic vulnerabilities and availability limitations, which rise the level of risks caused by their utilisation.

Numerous complex systems and services like

1. navigation at sea, in the air and on the roads and railways
2. Search and Rescue operations
3. emergency health services
4. crime control
5. telecommunications
6. traffic information services
7. fleet management etc.

Applications of local or national importance, as well as a plethora or personal services like personal navigation, location-based services etc. depend on the sustainable GPS performance.

Satellite signal availability and satellite visibility are among the sources affecting satellite navigation performance.

The ability of the US to maintain the GPS constellation has risen concerns on potential risks for numerous complex systems utilising satellite navigation.

The satellite outages can lead to the deterioration of the geometric dilution of precision (GDOP).

The number of visible satellites and their spatial distribution across the visible sky directly affect the dilution of precision (DOP).

One of the two components of the GPS positioning error is DOP.

GPS positioning error = UERE × DOP

where: UERE = User Equivalent Rage Error

DOP =. Dilution of Precision Dilution

The additional effect of a single GPS satellite outage on services that require more than a necessary number of visible satellites was observed.

As the time of at least four satellites visibility does not change significantly in case of a single GPS satellite outage, systems and services demanding a continuous visibility of more than the necessary four satellites may experience significant limitations.

GPS outage case-study revealed the measurable effects of a single GPS satellite outage on the DOP and GPS performance degradation, that potentially affects the performance of numerous systems relying on satellite navigation services.

A single GPS satellite outage causes both the DOP degradation and the extension of periods with a lower GPS satellites visibility.

Since the effects of the GPS performance degradation can be observed nation-wide, an impact on the national infrastructure and related actions in sustaining the system integrity should be considered (outage monitoring and notification, provision of assistance and augmentation to positioning etc.).

The effects of multi-satellite outages would be much more devastating than a single satellite outage.

END OF UNIT 3