SPATIAL GEODESY
LECTURE NOTES

JEAN-MATHIEU NOCQUET

GEOAZUR, UNIVERSITÉ CÔTE D'AZUR, INSTITUT DE RECHERCHE POUR LE DÉVELOPPEMENT (IRD), CNRS, OBSERVATOIRE DE LA CÔTE D'AZUR

INSTITUT DE PHYSIQUE DU GLOBE DE PARIS

Preliminary version - Last update November 2021

Contents

1. Introduction

1.1   See the invisible

Seeing the invisible. Many sciences share this goal, but Geodesy is among the most fascinating disciplines, because it allows us to see the deep reasons that explain why our terrestrial environment is how we see it. While geology aims at deciphering the past history of the Earth, with geodesy, we can see in "real-time" the processes shaping the Earth and try to understand them.

You probably have travelled, or perhaps hikes or climbed in the Alps or the Pyrenees. Do you know if the Alps are still growing today? Why are they where they are, squeezed between a the Mediterranean sea, a flat plain in northern Italy and France? Has anyone look at a satellite image of Asia? It clearly shows that Asia has been deformed by the indentation of India moving northward. Is India still pushing Asia? How is organized the wide deformation from the Himalaya to Siberia that we guess from the images? By providing millimeter per year measurements of surface velocities, geodesy can image the present-day evolution of actively deforming zone.

Every year, as your nails are growing of a few tens of millimeters, the Earth crust moves, but unlike the nails, it also deforms, and as a consequence stress increases. Earthquakes are the sudden and sometimes catastrophic release of the accumulated stress. While large earthquakes are visible, the processes leading to them has yet to be understood. Do you think that these slow processes are constant through space and time? Geodesy allows to see how the Earth is preparing the next earthquake, opening the perspective to anticipate them in future.

Beside the dynamics of the Solid Earth, GPS geodesy had problems with noise. Reducing the noise for solid Earth applications, led to some discoveries: GPS can detect magnetic perturbations in the upper atmosphere and is now used for space weather. Because the electromagnetic wave used by GPS is delayed when travelling across the lower atmosphere, GPS can be used to evaluate the amount of water vapor, providing new data for meteorological models. Finally, GPS also led to another unexpected contribution. Because oceans, atmosphere constantly move at the Earth surface, they deform it. Together with Spatial gravimetry, space geodesy gives us access to the large scale dynamics of fluid envelopes, and in particular to the continental scale hydrology.

1.2   Content and objectives of the lecture

By the end of this lecture, you should know the basics of GPS positioning and more generally of Global Navigation Satellite Systems (GNSS). You will have a flavor of the gory details required to move from the meter accuracy provided by your cell phone to the millimeter accuracy required for the scientific application.

You should be able to read and understand scientific papers presenting GPS data. You should have some methodological tools to extract the relevant information in terms of velocity, displacement and strain. You will also understand how models are built using GPS results as input data to make inferences about the processes.

1.3   More details

1.3.1   Contribution of space geodesy to the ordinary life

Although, this lecture will be focusing on Science, it is interesting to see how geodesy through satellite navigation systems have impacted our life. The most widely used application is probably through cell phone which all include a GPS (sometime GNSS) antenna and chip allowing to calculate your position in real time. Without knowing it, you are using one product of scientific geodesy. To get coordinates either on the Earth's surface or in space, you need a reference frame, that is an origin, three axis and a norm to express your coordinates. The GPS frame is called WGS84. Although determined by the DoD, WGS84 is as close as possible to the scientific International Terrestrial Reference Frame (ITRF) produced by the IGN-IPG Geodesy team (http://itrf.ensg.ign.fr). The availability of a consistent global reference frame has many implication. It allows a globally consistent (your coordinates system does depend on the country where you are) geo-referencing of aerial and satellite images, geographical information data bases and systems (GIS), maps, land surveying (cadaster), roads and civil engineering works (building a dam for instance).

1.3.2   Spatial geodesy as a global sensing tool of environmental changes

Because mass redistributions induce deformation and gravity changes, spatial geodesy can be used to monitor both changes in studies of:

- sea level
- changes in oceanic circulation
- circulation of continental water
- antarctic and arctic ice cap retreat

Because corrections on the electro-magnetic wave travel time are needed, GNSS data can be used to:

- sensing the water content of the atmosphere at local and global scale, for meteorological predictions or even climate studies
- evaluate the electro-magnetic activity of the ionosphere (space weather monitoring), which has become crucial for communication satellites

2. GPS principles

2.1 Introduction

GPS is now the most common geodetic technique used in geophysics. GPS relies on a constallation of satellites emitting electro-magnetic waves. It is based on the travel time measurement between satellites and a receiver.

Some preliminary remarks:

- Because GPS signal uses centimer wavelength (L-band, 1.4 Ghz 20 cm), it can work under any meteorological conditions
- Signal is emitted by the satellite: any users on the ground can benefit from the signal
- The use of the signal is free and publically available (but with some restriction)
- A GPS receiver with its antenna is only 3 kg. Measurements can be made almost everywhere
- Cost is moderate (5-10 k$)
- Accuracy: currently a few millimeters in a GLOBAL frame for daily position

NAVSTAR GPS (NAvigation System by Timing and Ranging - Global Positioning System) is first a project from the Department of Defense (DoD) of the USA. It started in 1973 and the first GPS satellite was launched in 1978. It became fully operational in 1995. GPS was created for military purposes, with the goal to provide precise (meter level) positioning anywhere on Earth or close to it, anytime, under any meteorological conditions, in real-time.

GPS evolved towards a more civil tool, encouraged by new directives from the American congress and some limitations were waived in 2000. Although designed for navigation and military purposes, geodesists found early in the 80s a way to use the signal differently in order to derive sub-centimetric coordinates. This accuracy however is at the price at specific equipment and relatively heavy post-processing further described in chapter **??**.

2.2 Principle of positioning

We measure the travel time of an electromagnetic wave between a satellite and a receiver. Since the signal is travelling at the speed of light, we know the distance between the satellite and the receiver.
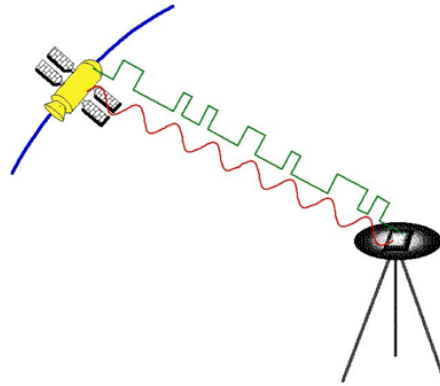
Figure 2.1: Schematic view of GPS signal: an electro-magnetic wave is sent by a GPS satellite. This wave is modulated so that it carries a binary code. While only code information are used in metric real-time positionning, high precision GPS uses both phase measurements and code.

The GPS constellation consists in 28 satellites orbiting at 20 000 km, along 3 plans having a 55.5^o inclination with respect to the equatorial plan. At 20 000 km, every satellite complete two full orbits every day. If you want to know how clever this choice is, you can go to (NASA-JPL GPS orbits movie). The main point is that any site of the earth at any time can receive signal from at least 4 satellites.

The signal sent by a GPS satellite is made of a code including the date of emission (t_e) and the position of the satellite. As a consequence, when we receive a signal at t_r , we know that we are somewhere on a sphere having the satellite as its center and of radius $r = c(t_r - t_e)$, where c is the speed of light (300 000 km/s). If we have a second satellite available, our position is at the intersection of two spheres (then on a circle). But, we want to know where we are, i.e. determine our coordinates (3 unknows). Then, 3 equations or observations are required. We therefore need 3 satellites to be able to determine a position. But finally, because there are clocks errors, we need a 4th satellite.

2.2.1 Pseudorange

Receiver clock errors dominate the measurements. Indeed, in order to be remain relatively cheap, GPS receivers have cheap clocks that are not accurate enough for positionning. If we assume a 1 millisecond error, this translates in $10^{-3} * 310^8$ m, that is 300 km. In satellite navigation, distance is the true geometric distance, range is the distance from satellite to receiver, plus path delays. Pseudorange is range plus clock errors.
Note: VLBI and GPS are pseudoranging techniques, SLR is ranging.
So a pseudorange is:

$$P = \text{true distance} + c * (\text{clock errors}) + c * (\text{path delays}) \qquad (2.1)$$

2.3 Orbits

Position of the emitting satellite are continuously broadcast by the satellite. Broadcast orbits have precision of 40 m, allowing positioning with a few meters (2-4 m) precision. Orbits are described by their 6 Keplerian elements that describe an elliptical trajectory:
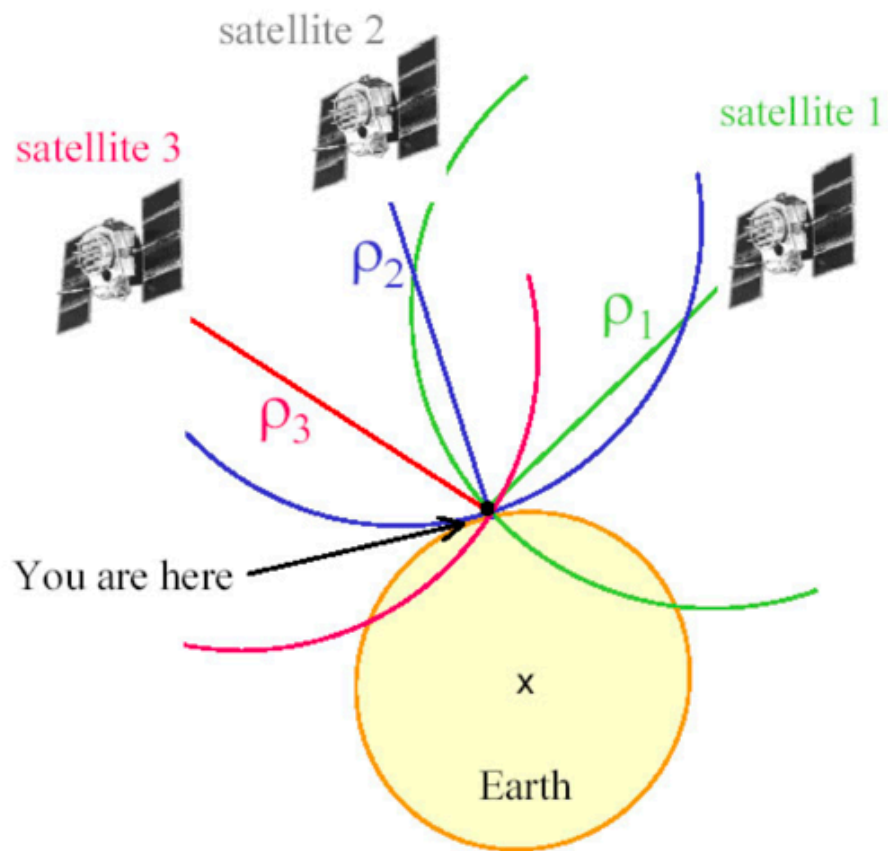
Figure 2.2: Schematic of positionning using trilateration
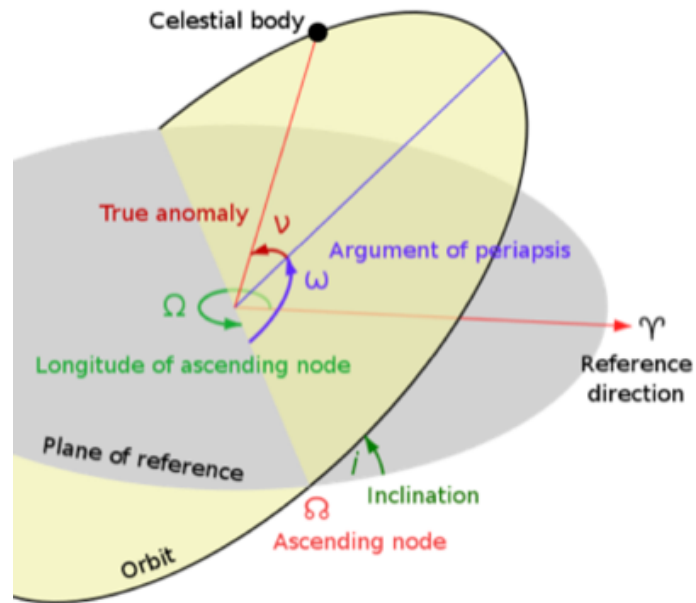
Figure 2.3: Keplerian elements describing an orbit.

- Semi-major axis of the ellipse
- Eccentricity of the ellipse
- Four angles (see Figure 2.3)

Earth is not a point mass, so satellite orbits are not exactly elliptical and broadcast message also include some corrections to the ellipse. Orbits are calculated by integrating the equations of motion (fundamental law of dynamics) for satellites from initial conditions (position, velocity at some time), subject to earths gravity field.

Since GPS navigation has to be real-time, orbits (and clocks) have to be predicted to be broadcast. DoD has setup a tracking network, with real-time telemetry. Orbits are therefore predicted adjusting past observations collected at sites with good approximate positions.

2.4   the GPS signal

2.4.1   How to code a signal?

The GPS positioning requires to transmit information (position of satellite and date of emission) from the satellite to the receiver. The signals from a GPS satellite are fundamentally driven by an atomic clocks (usually cesium, which has the best long-term stability). The fundamental frequency is 10.23 Mhz. Two carrier signals, which can be thought of as sine waves, are created from this signal by multiplying the frequency. The reason for having different wavelengths is for correcting the delay of the signal in the Earths ionosphere, as we will see later.
GPS uses two different waves:
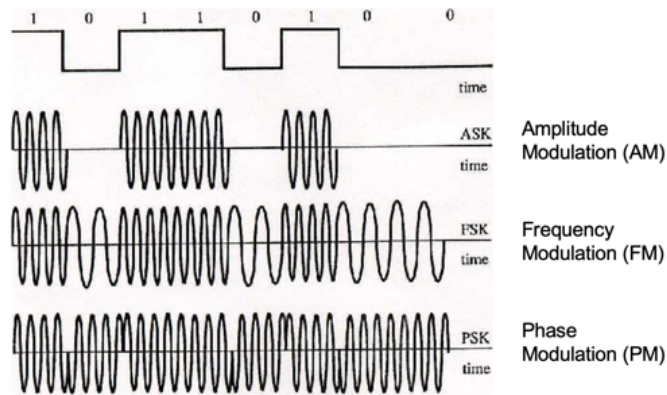
Figure 2.4: Different signal modulations used to carry information

- L1 at 154*10.23 MHz = 1.57 GHz (19 cm)
- L2 at 120*10.23 MHz = 1.22 GHz (24 cm)
- more recent satellites also have L5 at 115*10.23 MHz 1.17 GHz (25 cm)

The carrier electromagnetic wave is sent continuously by each satellite. This sinusoidal wave is modulated by a code, which is a series of 0 & 1. The wave is called carrier phase. There are different ways to "put" code on a carrier wave using modulation of the electro-magnetic wave. They are illustrated in Figure 2.4. GPS uses the phase modulation.

2.4.2 GPS code description

A code includes the following information:
- A number identifying the satellite called PRN
- The precise date of emission of the code
- The satellite position at the emission time
- Additional information: health of the satellite, corrections

GPS uses several codes. Codes are characterized by the way they can be decoded and their chip length which is the distance associated with each bit of the code.
- C/A (coarse acquisition) is made at 1.023 MHz on L1, sending 50 bits per second and has a chip length of 293 m. For a few years, a similar code is also sent on L2, called L2C.
- P1 : code P on L1 (29.3 m)
- P2 : code P on L2 (29.3 m)

The expected precision on the tag timing for the receiver is a fraction (a few %) of each code length.

2.4.3 Denial of accuracy

US DoD can reduce accuracy for real-time civilian users, through several techniques.

Selective Availability S/A

S/A has been on from 1990 to late 1990s. It introduces errors in navigation message (Epsilon), rapid fictitious variations in GPS satellite clocks (Dither). Military GPS receivers have special chips to handle these effects. The SA has been suppressed by the Clinton administration in May 2000 to encourage GPS civil applications and growth of the GPS "market".

Code Correlation for Ranging
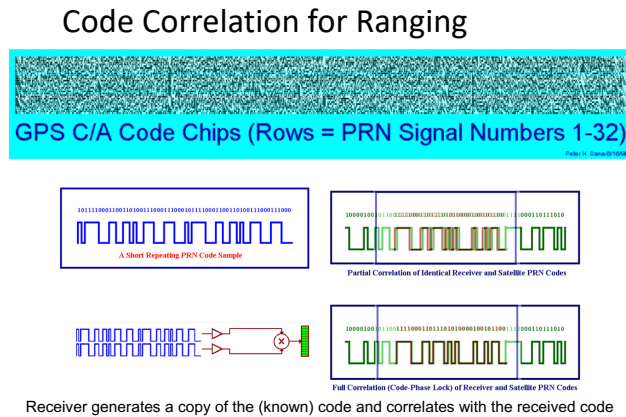
Figure 2.5: Receiver code correlation for ranging

Anti-spooging A/S

A/S is an an additional encryption (called Y-code) of the P-code to prevent the enemy from imitating (spoofing) GPS signal. It is on since 1994. Modern receivers get around this encryption in various ways (including a near-reverse-engineering in one case).

2.4.4 How do GPS receiver work?

In order to decipher the code, each receiver is running a continuous loop for each PRN. It then continuously makes a cross correlation of its internal signal with the signal received from the antenna. The signal is decoded when the cross-correlation coefficient is high (Figure 2.5). This implies that the receiver should know some sequence of the code before it receives the signal. Some code (or part of them) are public, some other have restrictions.

2.5 The pseudo-range solution

In this paragraph, we derive the equations to solve for the coordinates X,Y,Z in geocentric cartesian coordinates from pseudorange measurements.

For a single satellite and a receiver, the correlation time shift gives an estimate of the travel time:

$$\text{Travel time} = T - T^{(S)} \tag{2.2}$$

with
- T time of reception at the receiver
- $T^{(S)}$ time of emission at the satellite

The pseudorange $P^{(S)}$ is:

$$P^{(S)} = (T - T^{(S)})c \tag{2.3}$$

with
- T = receiver clock reading at reception

- $T^{(S)}$ = satellite clock reading at transmission
- c = speed of light = 299792458 m/s

Here, we assume that the satellite clock error is small, thanks to highly precise Cesium or Rubidium clocks on-board of satellites. Noting t^S and t the true time at the satellite and receiver respectively, we have $t^S = T^S$ and accounting for the receiver clock bias τ, $T = t + \tau$. We further make the approximation that the position sent in the code is exactly the position of the satellite at t^S (see http://www.nbmg.unr.edu/staff/pdfs/blewitt/basics of gps) for more details.

Then we have:

$$P^{(S)} = (t + \tau - t^{(S)})c = (t - t^{(S)})c + \tau c = \rho^{(S)} + \tau c \tag{2.4}$$

$(t - T^{(S)})c = \rho^{(S)}$ is the true distance between the satellite and the receiver, which can be expressed as a function of the receiver position $[x, y, z]$ at time t:

$$\rho^{(S)} = [(x^{(S)} - x)^2 + (y^{(S)} - y)^2 + (y^{(S)} - y)^2]^{1/2} \tag{2.5}$$

Now, generalize to multiple satellites $S = 1, 2, 3, 4$. We use a superscript (S) to identify each satellite S (dont confuse with an exponent). Later we will have to use a subscript to keep track of multiple receivers:

$$
\begin{aligned}
P^1 &= [(x^{(1)} - x)^2 + (y^{(1)} - y)^2 + (z^{(1)} - z)^2]^{1/2} + c\tau \\
P^2 &= [(x^{(2)} - x)^2 + (y^{(2)} - y)^2 + (z^{(2)} - z)^2]^{1/2} + c\tau \\
P^3 &= [(x^{(3)} - x)^2 + (y^{(3)} - y)^2 + (z^{(3)} - z)^2]^{1/2} + c\tau \\
P^4 &= [(x^{(4)} - x)^2 + (y^{(4)} - y)^2 + (z^{(4)} - z)^2]^{1/2} + c\tau
\end{aligned}
\tag{2.6}
$$

For fast solution, we want to solve this equation as a linear system (see linear the Appendix on linear algebra). We linearize our equations about approximate values (x_0, y_0, z_0, τ_0) and note $x = x_0 + \Delta x, y = y_0 + \Delta y, z = z_0 + \Delta z$. The Taylor series at degree 1 gives:

$$P^{(S)}(x, y, z, \tau) = P^{(S)}(x_0, y_0, z_0, \tau_0) + \frac{\partial P^{(S)}}{\partial x}\Delta x + \frac{\partial P^{(S)}}{\partial y}\Delta y + \frac{\partial P^{(S)}}{\partial z}\Delta z + \frac{\partial P^{(S)}}{\partial \tau}\Delta \tau + \varepsilon \tag{2.7}$$

which can be written as:

$$P^{(S)}(x, y, z, \tau) - P^{(S)}(x_0, y_0, z_0, \tau_0) \approx \begin{bmatrix} \frac{\partial P^{(S)}}{\partial x} & \frac{\partial P^{(S)}}{\partial y} & \frac{\partial P^{(S)}}{\partial z} & \frac{\partial P^{(S)}}{\partial \tau} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{bmatrix} + \varepsilon \tag{2.8}$$

So we can write a linear system $d = Gm$, by stacking the individual observation equations:

$$\begin{bmatrix} P^{(1)} - P^{(1)}(x_0, y_0, z_0, \tau_0) \\ P^{(2)} - P^{(2)}(x_0, y_0, z_0, \tau_0) \\ P^{(3)} - P^{(3)}(x_0, y_0, z_0, \tau_0) \\ P^{(4)} - P^{(4)}(x_0, y_0, z_0, \tau_0) \end{bmatrix} = \begin{bmatrix} \frac{\partial P^{(1)}}{\partial x} & \frac{\partial P^{(1)}}{\partial y} & \frac{\partial P^{(1)}}{\partial z} & \frac{\partial P^{(1)}}{\partial \tau} \\ \frac{\partial P^{(2)}}{\partial x} & \frac{\partial P^{(2)}}{\partial y} & \frac{\partial P^{(2)}}{\partial z} & \frac{\partial P^{(2)}}{\partial \tau} \\ \frac{\partial P^{(3)}}{\partial x} & \frac{\partial P^{(3)}}{\partial y} & \frac{\partial P^{(3)}}{\partial z} & \frac{\partial P^{(3)}}{\partial \tau} \\ \frac{\partial P^{(4)}}{\partial x} & \frac{\partial P^{(4)}}{\partial y} & \frac{\partial P^{(4)}}{\partial z} & \frac{\partial P^{(4)}}{\partial \tau} \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \\ \Delta \tau \end{bmatrix} \tag{2.9}$$

Let's evaluate the partial derivative to get G, using $\frac{d\sqrt{u}}{dx} = \frac{du/dx}{2\sqrt{u}}$.

$$
\begin{aligned}
\frac{\partial P^{(S)}}{\partial x} &= \frac{-2(x^{(S)}-x_0)}{2(x^{(S)}-x)^2+(y^{(S)}-y)^2+(z^{(S)}-z)^2]^{1/2}} &= \frac{x_0-x^{(S)}}{\rho^{(S)}} \\
\frac{\partial P^{(S)}}{\partial y} &= \frac{-2(y^{(S)}-y_0)}{2(x^{(S)}-x)^2+(y^{(S)}-y)^2+(z^{(S)}-z)^2]^{1/2}} &= \frac{y_0-y^{(S)}}{\rho^{(i)}} \\
\frac{\partial P^{(S)}}{\partial z} &= \frac{-2(z^{(S)}-z_0)}{2(x^{(S)}-x)^2+(y^{(S)}-y)^2+(z^{(S)}-z)^2]^{1/2}} &= \frac{z_0-y^{(S)}}{\rho^{(S)}} \\
\frac{\partial P^{(S)}}{\partial \tau} &= &= c
\end{aligned}
\tag{2.10}
$$

So,

$$
G = \begin{bmatrix}
\frac{x_0-x^{(1)}}{\rho^{(1)}} & \frac{y_0-y^{(1)}}{\rho^{(1)}} & \frac{z_0-z^{(1)}}{\rho^{(1)}} & c \\
\frac{x_0-x^{(2)}}{\rho^{(2)}} & \frac{y_0-y^{(2)}}{\rho^{(2)}} & \frac{z_0-z^{(2)}}{\rho^{(2)}} & c \\
\frac{x_0-x^{(3)}}{\rho^{(3)}} & \frac{y_0-y^{(3)}}{\rho^{(3)}} & \frac{z_0-z^{(3)}}{\rho^{(3)}} & c \\
\frac{x_0-x^{(4)}}{\rho^{(4)}} & \frac{y_0-y^{(4)}}{\rho^{(4)}} & \frac{z_0-z^{(4)}}{\rho^{(4)}} & c
\end{bmatrix}
\tag{2.11}
$$

So, we get

$$
\begin{bmatrix}
\Delta x \\
\Delta y \\
\Delta z \\
\Delta \tau
\end{bmatrix}
= G^{-1} d
\tag{2.12}
$$

This assumes that G^{-1} exists. It will exist as long as there are 4 or more satellites located in distinct directions in the sky. Two satellites located in exactly the same place would count as one. In general, we will have more than 4 satellites observed at a time. In that case of over-determined inverse problem, we find the best solution using Least squares. Least squares solution minimizes the sum of squares of residuals, that is $v^T v$ with $v = d - Gm$. The least squares solution, for equally weighted data, is

$$
\mathbf{m} = (G^T G)^{-1} G^T d
\tag{2.13}
$$

2.5.1 Position Dilution of Precision

The relative precision of horizontal and vertical components is determined by satellite geometry, which is implicity included in the A matrix, called the Design matrix.
Some remarks:

- No satellites below you, so vertical is intrinsically less precise
- Any path delay further reduces accuracy of the vertical coordinate.

Most handheld GPS reports a number called PDOP, which stands for Position Dilution of Precision, that are indicator of the precision. These numbers are derived from the variance-covariance matrix of $\mathbf{x}$. The variance-covariance matrix provides the information about the precision of estimated parameters together with their correlation. According to Gauss theorem

$$
C_{\mathbf{x}} = (A^T A)^{-1}
\tag{2.14}
$$

According our setup of the inverse problem, $C_{\mathbf{x}}$ in provided in XYZ geocentric cartesian coordinates. We can express it in the local frame by using the rotation matrix R from global gecentric frame to the local frame, by using the law of propagating covariance:

$$C_{\mathbf{x}}(ENU) = RC_{\mathbf{x}}(XYZ)R^T = \begin{bmatrix} \sigma_e^2 & \sigma_{en} & \sigma_{eu} \\ \sigma_{ne} & \sigma_n^2 & \sigma_{nu} \\ \sigma_{ue} & \sigma_{un} & \sigma_u^2 \end{bmatrix} \tag{2.15}$$

The different DOP give measures of how the satellite geometry maps into position.

$$\begin{aligned} VDOP &= & \sigma_u \\ HDOP &= & (\sigma_e^2 + \sigma_n^2)^{1/2} \\ PDOP &= & (\sigma_e^2 + \sigma_n^2 + \sigma_u^2)^{1/2} \end{aligned} \tag{2.16}$$

3. High precision GPS

3.1 From meter to millimeter

3.1.1 How scientists diverted the use of a military system

In the past chapter, we saw the basics for real-time positioning with a precision of a few meters. We saw that precision was limited by two characteristics of GPS:

- orbits broadcast in real-time have a precision of the order of 30-40 m
- pseudorange measurements (after receiver clock correction) have an intrinsic precision which is a fraction (5-10 percent) of the code length, which had length of 30 m for the most precise P-code.

In order to overcome these limitation, geodesists had two ideas. First, they found a way to create a very precise new observable from the GPS signal. Rather than using pseudo-ranges, they found that directly cross-correlating the code, cross-correlation of the phase of the carrier wave could be done at a precision equivalent to 0.3 mm! There are however some "price" to pay for that in terms of processing. Secondly, orbit determination had to be drastically improved. The next two chapters described those efforts.

3.1.2 The long way to millimeter precision

But improving the precision by 2 orders of magnitude also required many further improvements: path delays also have to be accounted for - or modeled - at the centimeter level. In observational sciences, one simple way to improve precision is to cumulate many observations so that the searched parameter precision improves with the square-root of the number of (independent) observations. So long observation sessions, typically of 24 hours during which a GPS antenna stays at the same place is one of the key ingredient. But, to determine a single X, Y, Z position, many factors are evolving through time. Think for instance of the Solid Earth tide which has an amplitude of 30 cm.

This chapter summarizes the different ingredients required to achieve a millimeter positioning.
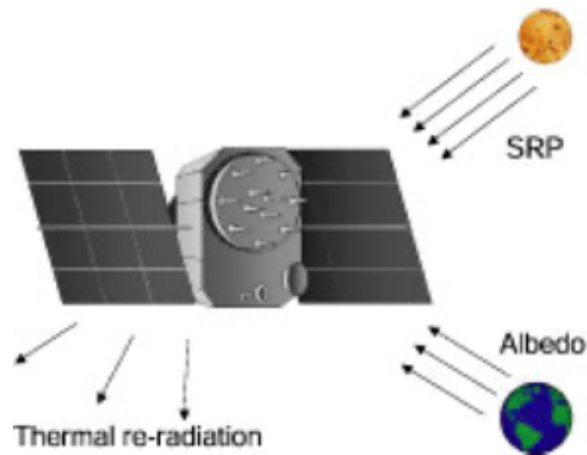
Figure 3.1: Radiations forces applying to a GPS satellite

3.2   High accurate orbits

3.2.1   Solar radiation pressure

Precise orbits are absolutely required for sub-centimeter positioning. Orbits are calculated by integrating the equations of motion (fundamental law of dynamics) for satellites from initial conditions (position, velocity at some time), subject to the forces applied to the satellite. Although the earths gravity field is dominant, other (non-conservative) forces also apply to the satellite. Particles emitted by the sun hit the satellite and transmit the motion quantity. This effect is known as as solar radiation pressure. Modelling of it requires to know how much area is normal to the sun and more sophisticated models account for the shape of the satellite. This effect is not negligible because GPS satellite have solar panel oriented towards the sun, to get their power. Of course, this effect vanishes as satellites are passing through the Earth'shadow. Aside the direct effect of solar radiation, the albedo of the Earth also contributes -to a smaller amount- of the radiation pressure applied to the satellites. Finally, there is a second-order effect of thermal re-radiation of the satellite.

3.2.2   The IGS global tracking network and orbit products

While integrating the equation of motion, orbits are also adjusted using observations made at a set of GPS stations, with well known coordinates, sampling the Earth'surface. Among the great successes of the geodetic community is the development of a federative network of GPS (and now GNSS) permanent, continuously recording network, which now includes more than 300 stations (Figure 3.2).

In practice, highly precise orbits are available from the International Service for GNSS (https://www.igscb.org/). There are 3 levels or precision:

- Ultra-Rapid: includes predict-ahead for real time use
- Rapid: Available next day
- Final: Available in <2 weeks

9 IGS Analysis Centers:

- 4 in USA: NGS, JPL/NASA, UC San Diego, MIT
- 1 in Canada (EMR)
- 4 in Europe: CODE, ESA, GFZ, GRGS

Figure 3.2: IGS tracking network. From http://igscb.igs.org
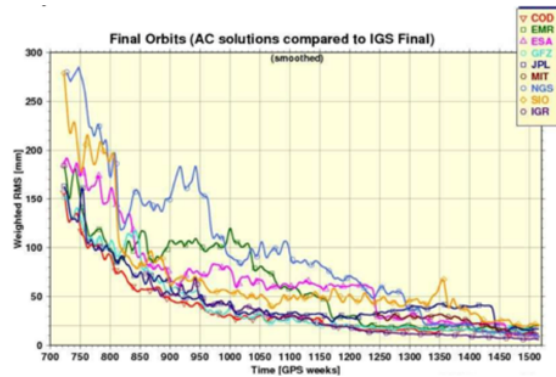
Figure 3.3: Evolution through time of the precision of IGS orbits. The figure shows the RMS (root mean square) compared to the final combined solution.

produce orbits that are combined and made publicly available by 4 IGS Global Data Centers, together with raw data of the tracking network, clocks, troposphere and ionosphere. IGN-France is one of the global data center (ftp://igs.ign.fr).

IGS orbits are more precise than broadcast ones because of three elements
- they use phase measurements in their adjustment
- they are adjusted using a wide network of 300-500 GPS fixed stations
- complex modelling of the orbit

The precision of orbits has improved through time. The precision for the final orbits used to be ~4.5 cm during the 1996-1999 period, but improved to ~1.3 cm in 2005-2007 (Figure 3.4).

3.3 The phase measurement

3.3.1 The phase observable

The principle of phase measurements can be easily understood through the analogy of a man with a precise clock sitting on a wharf and seeing the waves passing through him. So, a phase measurement in GPS is an initial value in $[-\Pi, \Pi[$ and then an integer number of cycles plus the phase on the current cycle. Although recent computers would have the speed required to follow the phase, high-precision GPS have a phase tracking system embedded in a dedicated chip.

Here, unlike the pseudo-range, we see that the observation equation relating the phase measurement to the satellite-receiver distance is ambiguous: we do not know the number of full cycles at

Figure 3.4: The phase measurement - analogy with a man sitting on a wharf with a clock. From Feigl and Cazenave, Formes et mouvements de la terre, Belin CNRS éditions, 1994.
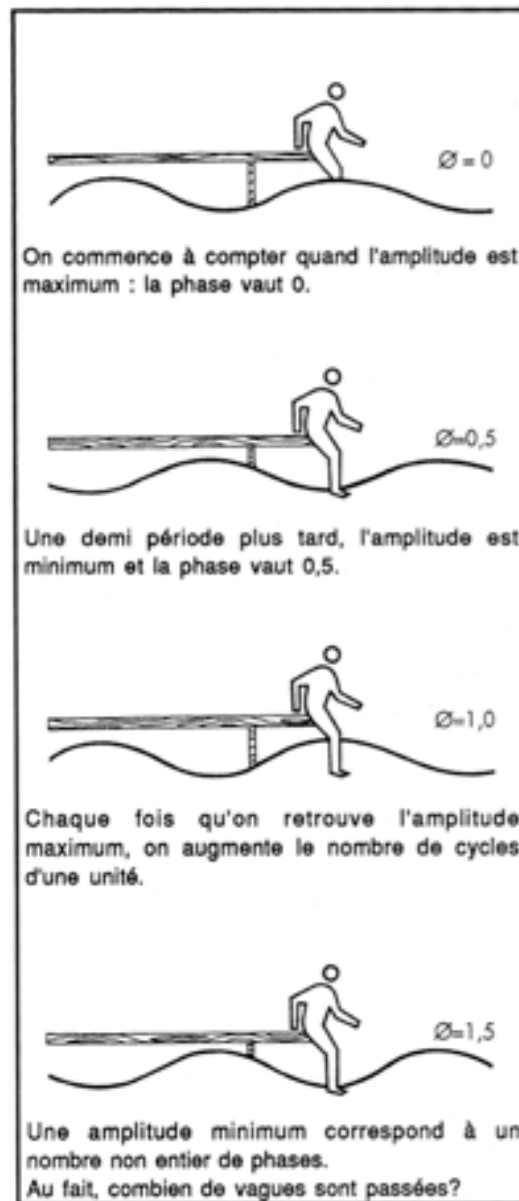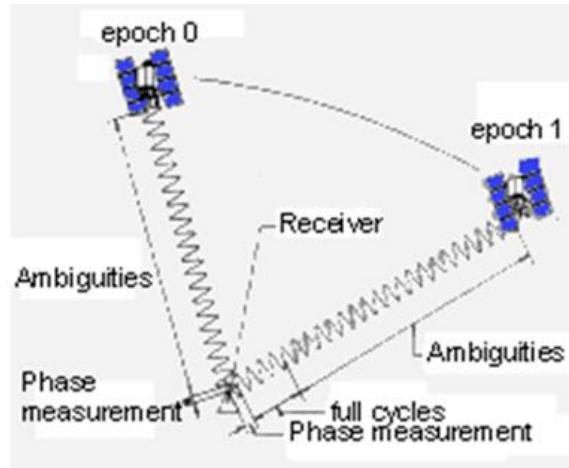
Figure 3.5: Phase measurement and the satellite-receiver distance

the first phase measurement. However, we know that this number is the same for all the phase measurements made by a receiver on the same satellite (Figure). So, an additional unknown parameter shows up in the equations. It can be resolved, simultaneously to the XYZ coordinates of the receiver if we have enough different measurements, therefore requiring long observations so that satellite position evolves sufficiently.

The relation between the satellite-receiver distance d is

$$d(t) = n.\lambda + \phi(t)\lambda = n.\lambda + \Phi(t) \tag{3.1}$$

where n is an integer number of unknown cycles called the "phase ambiguity" or simply "ambiguity" independent of t, λ the wavelength of the carrier (λ can be L1=19 cm or L2=24 cm) and ϕ the phase measurement (in radians) or $\Phi(t)$ converted in meters.

3.3.2 Double differences

As for pseudo-ranges, equation 3.3 is biased because of clocks errors. For precise positioning, the satellite clock error is also important and has to be corrected. The basic idea is that errors that degrade the precision are common (or highly similar) among two GPS receivers. We can therefore better determine relative coordinates than absolute.

Let's assume that we have a receiver A and B, together with satellite i and j. Starting with the phase measurements from satellite j, (dropping t), the phase observation equation is:

$$d_{A_j} = n_{A_j}\lambda + \Phi_{A_j} + \varepsilon_j + \varepsilon_A + \rho_{A_j} \tag{3.2}$$

with
- ε_j the clock error of satellite j
- ε_A the clock error of receiver A
- ρ_{A_j} the propagation delay

Similarly,

$$d_{B_j} = n_{B_j}.\lambda + \Phi_{B_j} + \varepsilon_j + \varepsilon_B + \rho_{B_j} \tag{3.3}$$
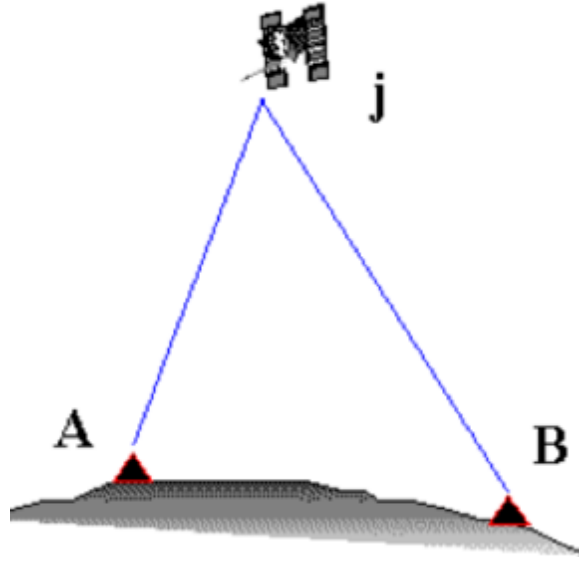
Figure 3.6: Simple phase difference

We form a new observation equation making the difference of the phase observation at *A* and *B*:

$$
\begin{aligned}
d_{A_j} - d_{B_j} &= (n_{A_j}\lambda + \Phi_{A_j} + \varepsilon_j + \varepsilon_A + \rho_{A_j}) - (n_{B_j}.\lambda + \Phi_{B_j} + \varepsilon_j + \varepsilon_B + \rho_{B_j}) \\
d_{A_j B_j} &= (n_{A_j}\lambda + \Phi_{A_j} + \varepsilon_A + \rho_{A_j}) - (n_{B_j}.\lambda + \Phi_{B_j} + \varepsilon_B + \rho_{B_j})
\end{aligned} \tag{3.4}
$$

ε_j, the satellite clock error has now disappeared. Moreover, if A and B are close, then $\rho_{A_j} \sim \rho_{B_j}$, so most of the propagation errors vanish. We can repeat this operation using a second satellite *i*, forming a double-difference:

$$
\begin{aligned}
dd_{ABij} &= (n_{A_j}\lambda + \Phi_{A_j} + \varepsilon_A + \rho_{A_j}) - (n_{B_j}.\lambda + \Phi_{B_j} + \varepsilon_B + \rho_{B_j}) \\
&\quad -(n_{A_i}\lambda + \Phi_{A_i} + \varepsilon_A + \rho_{A_i}) - (n_{B_i}\lambda + \Phi_{B_i} + \varepsilon_B + \rho_{B_i}) \\
&= (n_{A_j}\lambda + \Phi_{A_j} + \rho_{A_j}) - (n_{B_j}.\lambda + \Phi_{B_j} + \rho_{B_j}) - (n_{A_i}.\lambda + \Phi_{A_i} + \rho_{A_i}) - (n_{B_i}.\lambda + \Phi_{B_i} + \rho_{B_i})
\end{aligned} \tag{3.5}
$$

ε_A and ε_B, the receiver clock errors have now disappeared.

Most GPS processing software providing the highest accuracy use double differences. There are a number of remarks associated with the double-difference equation:

- The noise on double differences is higher than on individual observables (a few mm)
- The double-difference removes clock errors, make estimation faster (less parameters because clock error are not estimated)
- Phase bias parameters reduce to integer values, which is useful to help solving the inverse problem

Figure 3.7: Simple phase difference

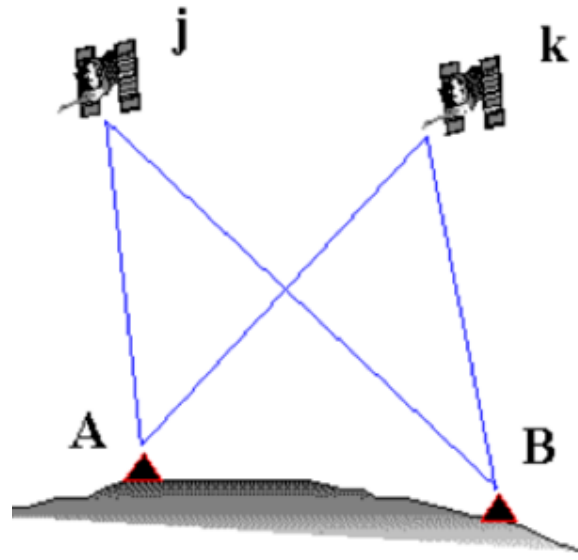- Requires a method to select differences, because you want only independent observations in the least-squares problem
- Requires to be careful when handling all indices when producing code

Finally, there is a very important consequence. When you difference between receivers, then you are actually now estimating the baseline vector between the two receivers, rather than the absolute positions of *A* and *B*. When processing a network, you actually end up with a set of coordinates where all relative positions are precise at the millimeter level, but the position of the network is not well defined in a reference frame. The obtained solutions are called "free" or "loose" solutions. They further need to be "attached" to an external reference frame. This point is described in XXXXX.

3.4 Ionosphere delay correction

So far, we were assuming that the GPS wave was travelling as a straight line at the speed of light. While this assumption is correct for space, the signal is delayed and curved when it enters the Earth atmosphere. The most important correction to the light speed is in the ionosphere.

3.4.1 The iono-free combination

Ionosphere is a layer between 60 and 1000 km above the surface where photons coming from the Sun have enough energy to separate electrons from the atoms and break molecules like O_2 into O ions. Depending on the solar activity (night/days or high activity periods of the solar cycle), the Total Electronic Content (or TEC) highly vary. More precisely, the TEC is the total number of electrons integrated between the satellite and the receiver, along a tube of one meter squared cross section. TEC unit is TECU=$10^{16} el/m^2$.

When crossing the ionosphere, the GPS wave will interact with the charged particles. Without entering too much into the details, the main effect of ionosphere is a delay of pseudo-range measurements, while the phase is advanced. Such delay can exceed 10 m and cannot be neglected, even

for meter precision navigation. Ionosphere is a dispersive medium at the frequency of GNSS (1-2 GHz). This is why GPS, and GNSS have several frequencies so that they can calibrate the delay in the ionosphere.

To a very good approximation, the path delay due to ionospheric refraction is proportional to $1/f^2$. Specifically, the path delay is $(40.3/f^2)$TEC, where TEC is the total electron content and for GPS $f = f1 = 1.57$ GHz for L1 and $f = f2 = 1.22$ GHz for L2. We can remove the effects of the ionosphere by forming a linear combination of the data at the two frequencies. Noting Φ_1 and Φ_2 the phase measurements on $L1$ and $L2$, we can produce an artificial phase observation ϕ_{LC} defined as:

$$\Phi_{LC} = -f1^2/(f2^2 - f1^2)\Phi_1 + f2^2/(f2^2 - f1^2)\Phi_2 \tag{3.6}$$

Indeed, for L1 and L2, the ionospheric biases ($\rho_{Ai}, \rho_{Aj}, \rho_{Bi}, \rho_{Bj}$ in the previous equations) are $(40.3\text{TEC}/f1^2, 40.3\text{TEC}/f2^2)$. Note that the two coefficients sum to 1. They have values of approximately (-1.54, 2.54). This removes 99.9% ionospheric effects and is called "iono-free". There are now second-order ionosphere models coming into use, which have an impact on positions at the few mm level or less.

Note that equation 3.6 also holds for the pseudo-range.

Coming back to navigation, most navigation GPS only record code on L1 (C/A and P codes, they are called single-frequency receiver) and cannot self-calibrate ionospheric delays. Actually, ionosphere is modeled in (almost) real-time and ionospheric corrections are broadcast in the navigation message so that user benefit from a first order correction.

3.4.2 Ionosphere dynamics using GPS

If one linear combination of the data can remove the ionosphere, then there must be an orthogonal combination that has nothing but the ionosphere:

$$\Phi_1 - \Phi_2 = -40.3TEC(1/f1^2 - 1/f2^2) + B_1 - B_2 \tag{3.7}$$

Here B_1 and B_2 are the phase ambiguities. This combination removes the unknown of position, satellite and receiver clock errors, as well as other sources for wave path delay. This combination, called "geometry-free" can be used to monitor the evolution of the ionosphere through time, or study ionosphere perturbation induced by earthquakes, tsunamis or volcanic explosions.

3.5 Troposphere delay correction

The atmosphere becomes denser as we get closer to the Earth surface. As a consequence, the GPS ray is bent and delayed, that is refracted. The refraction index controls the bending and the delay. The refraction index is a function of the gas composition and of the pressure.

The ratio Oxygen/Nitrogen is constant in the atmosphere. With a pressure model (a hydrostatic model is fine), we can therefore predict the delay induced by the dry air. However, the water vapor content of the troposphere is highly variable (especially in Ecuador!). This part is the most difficult to model.

The tropospheric delay is classically divided into:
- a hydrostatic contribution (order of magnitude 2.3 m)
- a water vapor contribution

Water vapor is the main factor currently limiting the precision of GPS, especially on the vertical component (1 cm). Troposheric delay corrections are estimated together with the coordinates during the processing. Troposheric delays change through time according to the line-of-site (azimut and elevation) of the satellite seen from the receiver. Estimating for each measurements in all the directions would make the problem very undetermined. In order to reduce the number of parameters, a single zenital parameter (at the vertical of the receiver) is solved every few hours (typically 2 hours). The troposheric path delay at some elevation angle e is $ZTD * mf(e)$. The function mf is called mapping function, mapping a value at elevation e from its value at zenith.

There are several mapping functions developped by the geodetic community.

- VMF mapping functions vary with space and time based on distribution of water vapor from meteorological models.
- GMF is seasonal average of VMF

The estimated values at the zenith are called ZTD for Zenital Total Delay. ZTD can be converted to water vapor mass also called Precipitable Water Vapor (PWV). This information is nowdays used for meteorological and climate studies.

3.5.1 Other corrections

Small motions during fixed measurements

We saw that improved precision requires long observation session, typically 24 hours long, to average errors. During a day, however, there is a deformation of the Solid Earth, that needs to be accounted for to estimate fiw coordinates. These effects include:

- Solid Earth tides in response to the Sun and Moon gravitatioanl attraction
- Solid earth responds to changing load of ocean tides. Displacements can be as large as 15 cm over 12 hours in Brittany for instance, where tidal range is large. Accurate removal depends on good tidal models

3.5.2 Phase Center Variation (PCV)

So far, we talk about receiver. For accurate positioning, a called geodetic antenna is required. A geodetic antenna allows first to record the L2 signal which has less energy than L1, and optionally removes reflected signals coming from the ground. The antena includes an amplifyer of the signal transmitted to be sent the receiver for decoding. Basically, an antenna is a solenoid, while all equations have been written as phase was measure at a stable fix ideal point in space. In the real world, the is an apparent delay or negative delay changing with the elevation (and azimut) of the satellite. This delay is different for L1 and L2.

Antenas are calibrated by the geodetic communuty.

3.6 Solving the inverse problem

3.7 Linear combination of observables

Compared to some classical problems, estimating coordinates is performed using least-sqaures but with the information that some unknown parameters are integer. As a consequence, most approaches first try to solve for the ambiguities through iterations before proceeding to the least-sqaures estimation. A difficulty is that λ is of the order of 20 cm. So, solving for ambiguities would require to first get an approximate solution of the order of a few centimters so that an ambiguity

having a float value of n x λ + 1.5 cm could be fixed to $n\lambda$. Pseudo-range solutions, even double-differenced and averaged over one day can only lead to a precision of a few tens of cm. So, the trick is to consider other artificial linear combinations to solve the ambiguity and then convert the obtained ambiguities to their equivalent values for L1 and L2. This combinations are:

- Widelane: f1 - f2 has wavelength 86 cm, which is first solved
- Narrowlane: f1 + f2 has wavelength 10 cm

3.8   Solving the inverse problem

Without entering into the details, the scheme to obtain the unknown parameters (coordinates + troposheric delays) is:

- make a prediction of phase and pseudo-ranges from approximate coordinates and atmospheric models
- Estimate a float solution from the LC double-differenced LC combination using least-squares
- Resolve widelane ambiguities using Pseudorange data
- Use fixed widelane bias and ionosphere-free bias estimate
- Solve the narrow-lane ambiguities
- Rewrite the above equation in terms of the widelane/narrow lane ambiguity
- Perform the final adjustment

3.9   Other strategies

So far, we have considered the problem of high-precision positioning through long sessions and double-differences of phase. This approach is the most common in the geodetic community, but other strategies have been developped, that I briefly summarize.

3.9.1   Precise Point Positioning (PPP)

One problem with the double-difference approach is that the dimension of the inverse problem scales with the number of sites, and the processing time with the square of the number of sites. Precise Point Positioning gets around this limitation using the following strategy.

- first requires a double-differenced solution at the global scale
- from this solution, calculate precise estimates of the satellite clock error
- then, rather than differencing, estimate clock error together with station coordinates

The main advantage of this approach is that a single site can be processed independently of the others, preventing that some analysis center provide satellite clocks and orbits.

3.9.2   Rapid Static

Long session help to achieve the highest precision. With shorter sessions, precisionn decreases because there is no night/day average of path delay mismodelling errors and it is harder to fix ambiguities. Nonetheless, centimer level precision can still be achieved with short sessions. Typically, surveyor use 15 mn long session and move to the next point.

3.9.3   High rate kinematic GPS

Kinematic GPS would use the same strategy of double-difference except that the position is estimated at every epoch of measurements. This processing then has a lot of unknowns. Although millimeter precision remains out of reach, centimeter precision can be obtained.

One of the most impressive result of kinematic GPS is to obtain the dynalic displacement during earthquakes.

3.10 Reference frame

Because the most accurate positioning is obtained using double-differences of phase, free solutions are produced where part of the information about the "absoluteăž position is lost during the data reduction

A second step in the processing is required to express coordinates (position and velocity) in a reference frame, which should be stable through time

This procedure is known as reference definition or stabilization

It consists in estimating a translation, a small rotation and a scale factor (and possibly their rate of change) with respect to an external reference frame. This transformation is known as Helmert transformation

The external reference frame is usually the International Terrestrial Reference Frame (ITRF) which is updated every 3-4 years

The ITRF is derived from a combination of VLBI (scale), SLR (origin) and GPS global solutions. The orientation is conventional. The orientation rate is taken as a No-Net-Rotation condition (NNR). The combination is made at IGN-France

The physical meaning of the global (horizontal) NNR-frame is not clear. It is significantly different from the Hot Spot frame (up to 4-5 cm/yr)

3.11 GNSS

GNSS is the Global Navigation Satellite System

It already includes GPS and its enhancement (additional code) GLONASS : russian system SBAS (Satellite Based Augmentation System) is a system that supports wide-area or regional augmentation through the use of additional satellite-broadcast messages.

It will include:
- European Galileo
- Chine COMPASS
- Indian IRNSS
- Japanese QZSS

All systems should be inter-operable

3.11.1 CGPS and survey-mode measurements

GPS measurements for tectonic can be used in two modes: survey mode or continuous GPS (CGPS). In survey-mode, a small marker in anchored in the rock. A system ensure that the antenna is set up at the vertical of the marker. The marker is observed from 8 to 48 hours.
- Survey-mode is
- less expensive
- does not require continuous power
- does not require security
- enables dense network (50 sites)

Continuous GPS are permanent stations with power autonomy and most often with telemtry to transmit their data. A pillar or a permanent fixation system is used. CGPS requires:

- Requires
- power
- Security
- Provides the highest accuracy
- Provides daily solutions
- Enables error analysis

The signal is much richer (see the time series)

3.11.2 Strength and limitations of GNSS Geodesy

3.12 Summary

Geodesy provides a quantitative description of motion at the surface of the Earth.

- Consistent over spatial scales: from meters to global scale
- Time scales: a few Hz (seismic waves) to a few decades
- Precision: mm for position, a few mm/yr for rates
- 3D displacements, although the vertical component is less precise

2010　　2012　　2014　　2016　　2018　　2020

---
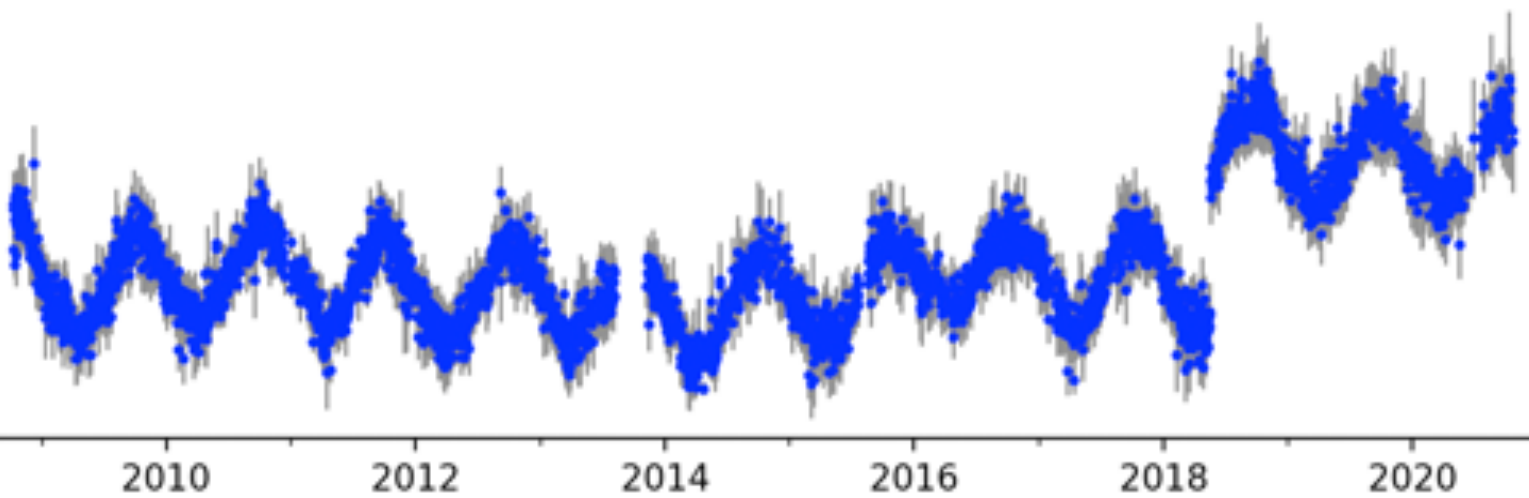
4. Time series analysis

4.1 Forewords and warnings

4.1.1 A bit of terminology

Here are defined a few words to avoid any confusion.

- Displacement is the change in position from one epoch to another epoch. It is expressed in m, cm, or mm.
- Velocity is the displacement rate. Ideally, it would be the time derivative of displacement and would depend on time. Most often in geodynamics, velocity is the average velocity over a period of time. Velocity are in m/s, but most often mm/yr are used.
- Deformation or strain is the relative displacement between two points divided by their relative distance. It has therefore no unit, but microstrain (that is a relative displacement of 1 mm over a distance of 1 km) or nanostrain (1 mm over 1000 km) are often used.
- strain rate is relative velocity between two points divided by their relative distance. It is there in s^{-1} or yr^{-1}, but nstrain/yr is commonly used.

Actually, we will see in the next chapter that the simple definition of strain and strain rate provided here is incorrect and requires more analysis.

4.1.2 Reference frame

This might be evident, but coordinates (position) and velocities are expressed in a reference frame. Changing the reference frame changes the values. A site showing Estward in a reference frame might be moving westward or even northward in another reference frame. Any interpretation therefore requires to have made clear before what is the reference frame.

4.1.3 Apparent and true motion

Coordinates are the results of a data analysis involving many models as we saw it in the previous chapter. As a result, coordinates changes are not always a consequence of a true motion induced by a physical process. For instance, changing an antenna on a CGPS site most often results in a jump of the vertical component. This is likely due to mismodelling of the antenna phase center for the
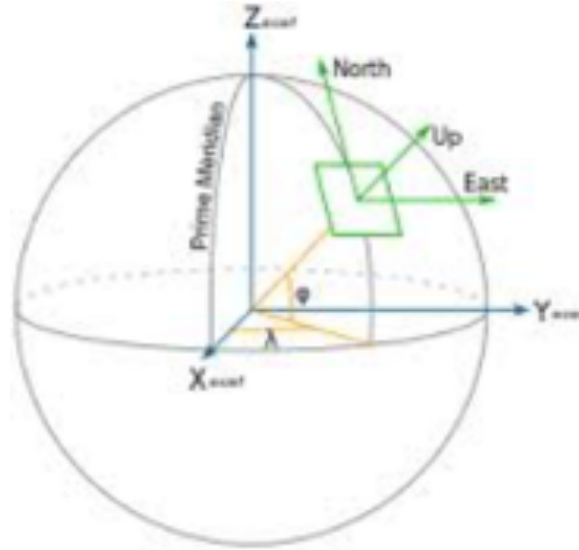
Figure 4.1: Transformation from geocentric cartesian to local East,North,UP coordinates

old, the new or both old and new antenna. Similarly, changing the orbits, the reference frame or tropospheric modeling results most often in jumps and sometimes trend in the time series.

As a consequence, we can be sure that part of the coordinate changes that we see are not caused by true motion, and part of the work, when trying to extract tiny signals is to discuss the uncertainties and ensure that true signal are greater than noise. Nonetheless, one good thing is that tectonic plates or Glacial Isostatique Adjustment motions cumulate through time. Expecting a few year, we can expect to increase the signal-to-noise ratio.

4.2   First description of GPS time series

Let's start looking at a first GPS time series. I chose the site BRAZ (Brazilia), located in the middle of the South America tectonic plate, within the Brazilian craton where neither earthquake nor slow deformation is expected. As a result of processing, coordinates are expressed in the International Terrestrial Reference Frame (ITRF). The ITRF is updated every few years or so, and the current realization of the ITRF is the ITRF2020. Raw coordinates are therefore expressed in the XYZ geocentric cartesian coordinates. But it is more convenient to see how the site moves in the North, East and UP components. So, to obtain time series in local coordinates, we first select a reference position X0,Y0,Z0 for BRAZ and define a local plan which is tangent to the ellipsoid and define in this plan, the East, North, Up axis, with X0,Y0,Z0 being the origin. Practically, we make use of the Rotation matrix from Global to Local coordinates R:

$$R = R_{gen \rightarrow loc} = \begin{bmatrix} -\sin\lambda & \cos\lambda & 0 \\ -\cos\lambda.\sin\phi & -\sin\lambda.\sin\phi & \cos\phi \\ \cos\lambda.\cos\phi & \sin\lambda.\cos\phi & \sin\phi \end{bmatrix} \tag{4.1}$$

with λ and ϕ are the longitude and latitude of BRAZ.

Figure 4.2 shows the obtained time series, with a blue dot corresponding to a position at each day. For the horizontal components, a trend is clearly seen. This trend is the secular velocity
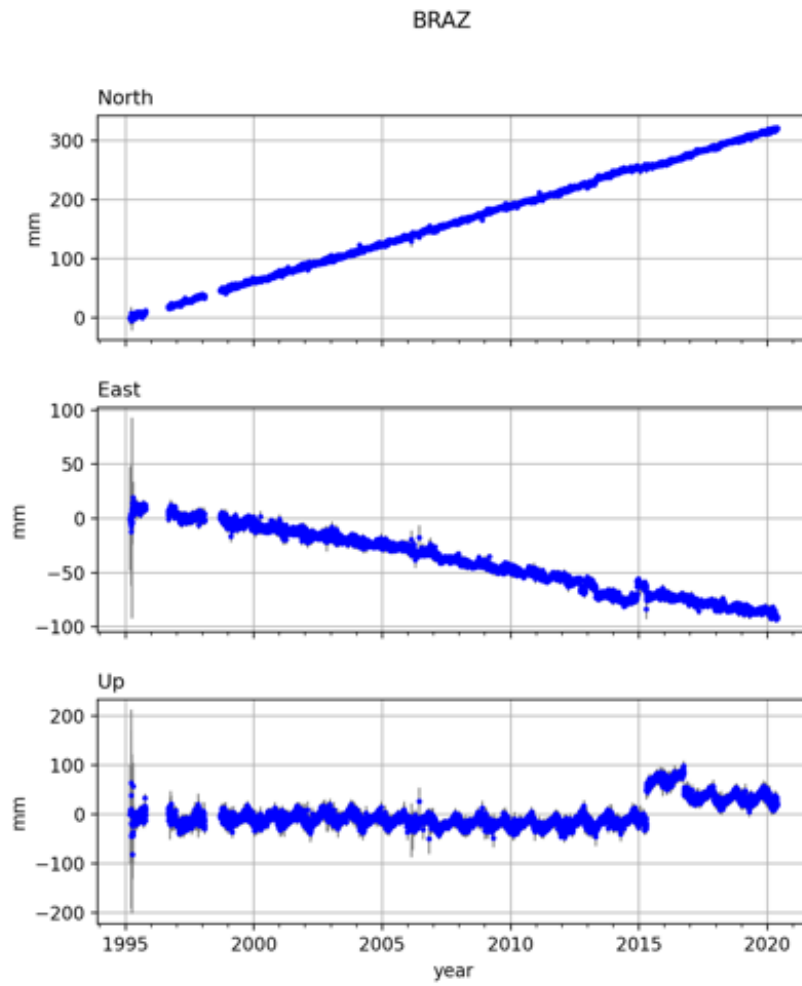
Figure 4.2: Raw time series in the ITRF2020 for IGS GPS site BRAZ (Brazilia)
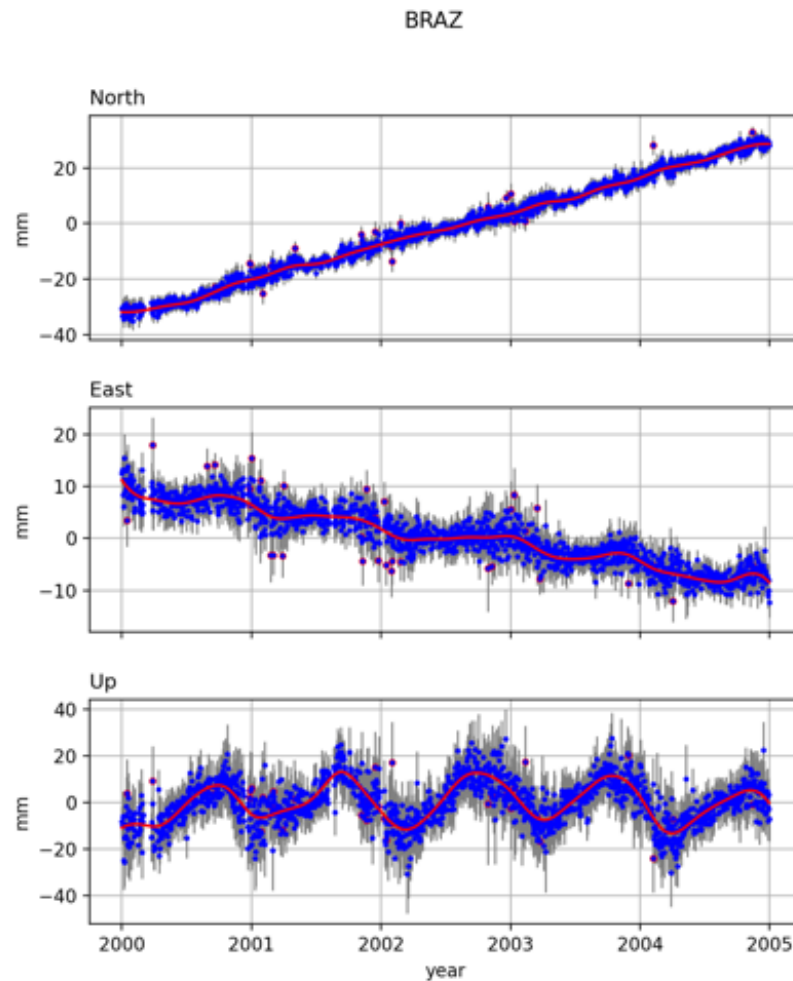
Figure 4.3: Raw time series in the ITRF2020 for IGS GPS site BRAZ (Brazilia) together with the low-pass filter time series (red) and outliers detection.

and in the case of BRAZ corresponds to the motion of the South America plate in the ITRF2020. We also see that there is a day to day scattering of the coordinates, that is high frequency noise. Each day comes with an uncertainty. This uncertainty, also called formal error comes from the covariance estimated at the end of a daily processing. It is called formal because it only depends on the adjustment and reflect the design of the least-squares problem when estimating the coordinates. We see that in 1995, there are large uncertainties up to 5-10 cm. We also can guess that there are some oscillations of the position, more clearly seen on the vertical component, superimposed on the trend. Finally, there is at least one obvious jump in 2015, due to an antenna change.

4.3 Basic trajectory model

Figure shows a zoom on the 2000-2005 period with a low-pass filtering. The low pass filter also helps identifying outliers, that are points despite having a good formal error depart from their neighbours.The low-pass filter also highlights that the oscillations have at the first order an annual period. So, a first trajectory model will include a position at a reference time, a slope and annual and semi-annual terms. Semi-annual terms are useful because they enable some asymmetry in the periodic annual signal. So, for a component $x(t)$, we write:

$$x(t) = x(t_0) + v(t - t_0) + a_{yr}\cos\left(2\pi t + \phi_{yr}\right) + a_{syr}\cos\left(4\pi t + \phi_{syr}\right) + v(t) \qquad (4.2)$$

where v(t) is the remaining error. This is a linear equation (as long as we develop the cos), that can be solved by linear least-squares.

The resulting time series once trend and seasonal terms have been removed is called a detrended time series. It is actually the residuals of the estimation. The long-term repeatability is often given to estimate the precision of the time series. It is the weighted root mean square of the residuals:

$$wrms_i = \sqrt{\frac{\sum v_i(t)^2/\sigma^2_{x_i(t)}}{\sum 1/\sigma^2_{x_i(t)}}} \qquad (4.3)$$

For BRAZ, we find:

- $wrms_N$ = 2.7 mm
- $wrms_E$ = 3.1 mm
- $wrms_U$ = 10.8 mm

The best component is the North, with the East being only slightly less precise. The vertical component is most often worst than the horizontal component by a factor of 3 to 5.

In terms of the parameters, we find:

```
###############################################################################
#  VELOCITY (mm/yr)
###############################################################################
    V_North      V_East       V_Up      S_V_North    S_V_East      S_V_Up
     12.53        -3.74       -0.46         0.04         0.05         0.14
###############################################################################
#  SEASONAL TERMS
###############################################################################
# Annual and semi-annual terms - phase in years
N_amplitude    N_phase   E_amplitude    E_phase    U_amplitude     U_phase
     0.4        0.1348       1.2          0.4020        9.5         -0.4933  (mm annual)
     0.5       -0.2455       0.4          0.4529        0.8          0.4909  (mm semi-annual)
```

Least-squares find that after 5 years of daily measurements, the velocity is determined at the 0.05 mm/yr! This is of course unrealistic, but why?

4.4 Colored noise

Least-squares in their most simple formulation assume that observation are independent, that is that the data covariance is null for off-diagonal terms. Then, uncertainty of velocity increases as the square root of the number of observations.
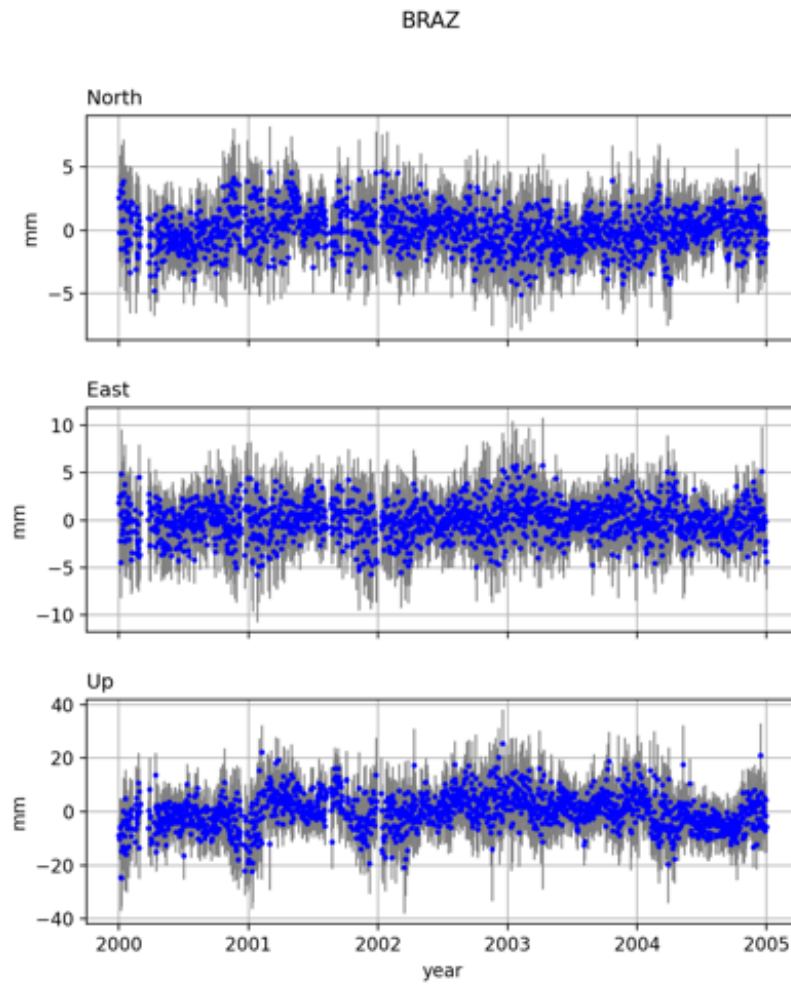
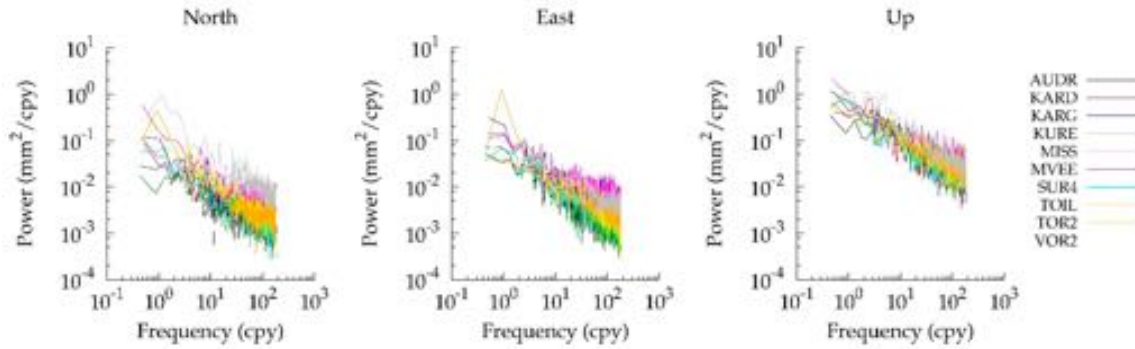Figure 4.4: BRAZ detrended time series

Figure 4.5: PSD plot for residuals GPS time series

One simple way to verify the assumption of no time correlation is a power spectrum density plot. White noise, that is pure random noise with no correlation has a flat horizontal spectrum. Colored noise are characterized by larger power at lower frequency.

All GPS time series have a PSD like the one shown in Figure indicating that some colored (that is time-correlated) noise dominates the time series.

La forme du spectre de puissance indique que le bruit peut être approximé par une loi en puissance du type:

$$P(f) = P_0 (\frac{f}{f_0})^\kappa \qquad (4.4)$$

où f est la fréquence, f_0 la fréquence à partir de laquelle on observe le bruit corrélé, P_0 la puissance spectrale à la fréquence f_0 et κ l'indice spectral traduisant l'augmentation de la puissance en fonction de la fréquence. κ est a priori un nombre réel, qui pour la plupart des processus physiques vaut entre -3 et -1. $\kappa = 0$ correspond au bruit blanc, $\kappa = -1$ au bruit de scientillation (flicker noise) et $\kappa = -2$ correspond à une marche aléatoire.

Une régression linéaire sur le spectre fournit une première approximation de f_c et de l'indice κ. Pour la plupart des stations et des analyses, on trouve que l'on quitte le domaine du bruit blanc autour de quelques semaines à quelques mois (< 3 mois). κ est en général proche de 1 (bruit de scintillation) pour les composantes horizontales, et des valeurs plus variées allant jusqu'au bruit de marche aléatoire pour la composante verticale.

Son ordre de grandeur est 2-5, 2-10, 7-23 $mm/yr^{1/4}$ pour les réseaux globaux et 3-5, 2-5, 7-12 $mm/yr^{1/4}$ pour les réseaux régionaux

Par ailleurs, les analyses spectrales ont mis en évidence une autre fréquence caractéristique des séries temporelles. En plus des termes annuels et semi-annuels, une analyse fine montre un pic à 1.040 cycles/année (351.2 jours appelée "année draconitique") présent dans les séries temporelles. Cette période correspond à la période de répétition de la géométrie de la constellation GPS dans son repère inertiel par rapport au soleil. Aucun pic à cette fréquence n'est détecté dans les séries temporelles VLBI et SLR suggérant donc qu'elle provient d'une erreur systématique propre au GPS, comme un défaut de modélisation des orbites ou d'effet du multi-trajet.

4.4.1 Origine du bruit corrélé dans les séries temporelles

L'origine du bruit corrélé a tout d'abord été recherchée dans la stabilité de la monumentation. Mais, le bruit corrélé est très fortement diminué dans le cas de très courte ligne de base (< 1 km). Ca ne de-

vrait pas être le cas si ce bruit provenait du mouvement du monument supportant l'antenne. De plus, il n'existe pas de corrélation claire entre les variations de la série temporelle et les paramètres environnementaux comme la température et les précipitations pouvant induire une dilatation/contraction du monument et la compaction du sol. A l'exception de quelques sites, il semble donc que les mouvements locaux ne contribuent que de manière marginale au bruit corrélé observé.

Un autre candidat est un défaut de modélisation des orbites. En particulier, des collègues en Allemagne ont montré que l'erreur de modélisation des antennes des satellites GPS se propagent dans les coordonnées, en particulier parce que la constellation GPS évoluent au cours du temps, avec l'arrivée et la retraite de certains satellites. De manière similaire, la stabilité long-terme des centres de phase des antennes au sol est questionnable. Cette hypothèse est renforcée par l'observation de la différence observée entre les calibration en site naturel et en chambre anéchoïque. L'environnement naturel évolue au cours des années (croissance des arbres, constructions de nouveaux batiments) qui peuvent affecter le comportement de l'antenne. Enfin, le dernier point peut être liés aux effets de surcharge que nous détaillerons plus loin.

4.4.2   Implication du bruit corrélé sur la détermination des vitesses

L'implication des caractéristiques des séries temporelles pour l'interprétation en terme de processus géophysiques géophysique est importante. la vitesse. Etant donné l'amplitude des termes saisonniers, typiquement de 2 mm pour les composantes horizontales et 4 mm pour la composante verticale, les études analytiques et paramétriques aboutissent à la recommandation qu'un minimum de 2.5 années est requis avant la publication de n'importe quelle vitesse. L'influence des termes périodiques devient négligeables à partir de 4.5 années d'observation.

La prise en compte des caractéristiques du bruit est plus problématique. Pour déterminer l'incertitude, il faut construire une matrice de covariance des positions à partir des caractéristiques du bruit. Cette matrice de covariance est supposée connue à un facteur près qui peut être estimé simultanément aux vitesses par des méthodes de maximum de vraisemblance. Une approche empirique simple est d'utiliser les résultats d'étude complètes qui montrent qu'en multipliant l'uncertitude (écart-type) sur les vitesses obtenue par moindres-carrés par un facteur compris entre 4 et 10, on reproduit a peu près les incertitudes de modèles plus complets.

4.5   Space correlated noise

GPS time series not only appear to be correlated in time, but also in space. Figure 4.6 shows detrended time series (seasonal terms not removed) in Brazil. They show similar variations at long period, but also at the period of a few days. Such common motion, real or not is called "common mode regional" motion. A way to quantify it is to calculate the correlation coefficient between two time series. Figure 4.7 shows this correlation as function of the distance separating two GPS sites. It clearly shows a positive correlation which progressively decreases with increasing distances, but remaining significant over distances as large as 1000 km.

When studying tectonic processes, such motion are unikely to have a tectonic origin and most studies would try to filter it. The lmost simple approaches consists in stacking the residuals time series to calculate a common mode. This common mode is then substracted to the original time series.
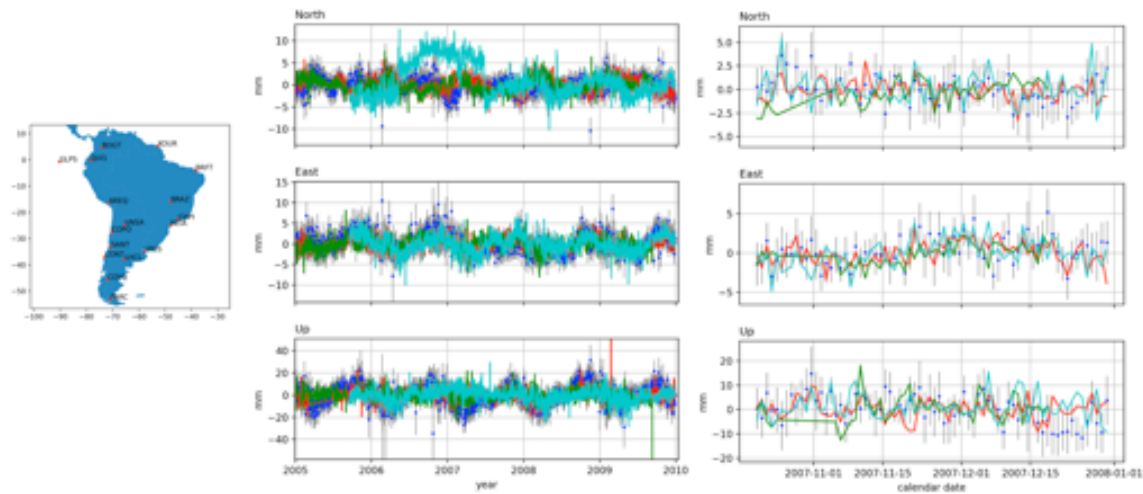
Figure 4.6: detrended time series for GPS BRAZ, NEIA, BRFT sites

Figure 4.8 shows the time series after regional common mode filtering. It shows that the daily scattering has reduced as well as long periods fluctuations.

4.6 Causes géophysiques du bruit

La comparaison de techniques indépendantes, par exemple SLR et GPS, ou GPS et Grace, montre que la composante annuelle est cohérente entre les techniques. Cette composante reflète donc un mouvement réel des stations.

Les enveloppes superficielles de la Terre (Atmosphère, Océan, Hydrologie continentale) sont en constante évolution et voient leur masse redistribuée à la surface de la Terre solide. Pour des fréquences de quelques minutes à quelques années, la Terre solide répond de manière élastique aux distributions de forces surfaciques qui lui sont imposées. Avec lavènement des mesures continues de pression, températures, altimétrie des océans, marégraphie, mesure des quantité deau contenues dans les premiers mètres du sol, il est aujourdhui possible dobtenir une carte globale des différents effets de charge.

Ces effets de charges ont une influence qui se sépare en deux composantes : (1) la Terre réagit de manière élastique à laccroissement de force surface appliquée sur la croûte (2) en se déformation, la Terre solide modifie champ de gravité et donc la force de gravité qui lui est appliquée. Elle se déforme donc aussi sous cette action. Les missions de gravimétrie spatiale ultra-précises (CHAMP, GRACE) sont donc incontournable pour étudier ces effets à grandes longueurs donde.

Les effets de charge sont calculés en convoluant la réponse élastique de la Terre à un dirac par la distribution surfacique de charge.

4.6.1 Surcharge hydrologique

La surcharge hydrologique est liée au stockage de leau sur les continents à travers les calottes de neiges sur les massifs montagneux, leau des réservoirs (lacs), leau contenue dans les aquifères et les premiers mètres du sol.
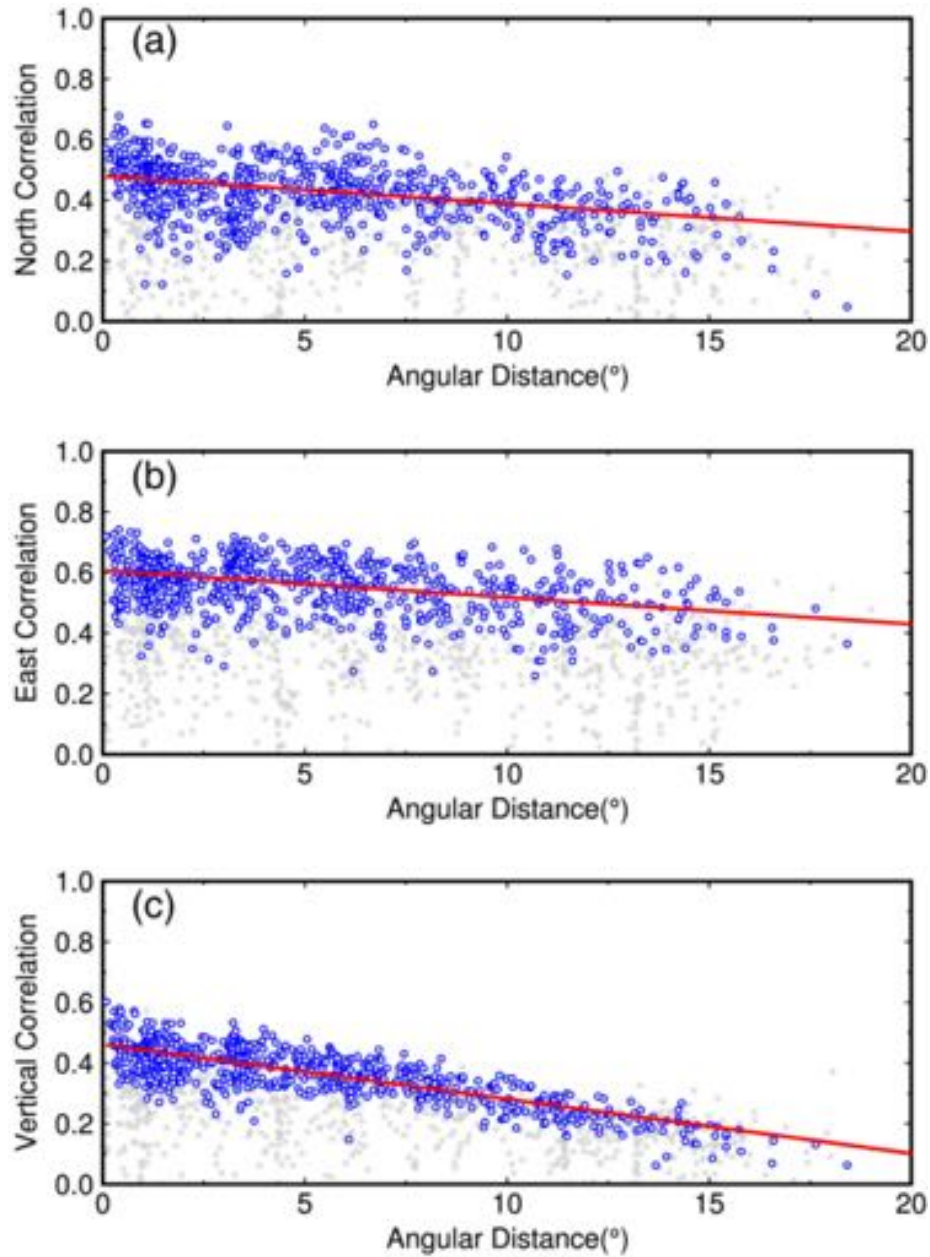
Figure 4.7: Correlation coefficient between pairs of GPS time series as a function of their relative distance. From Tian and Chen (2016)
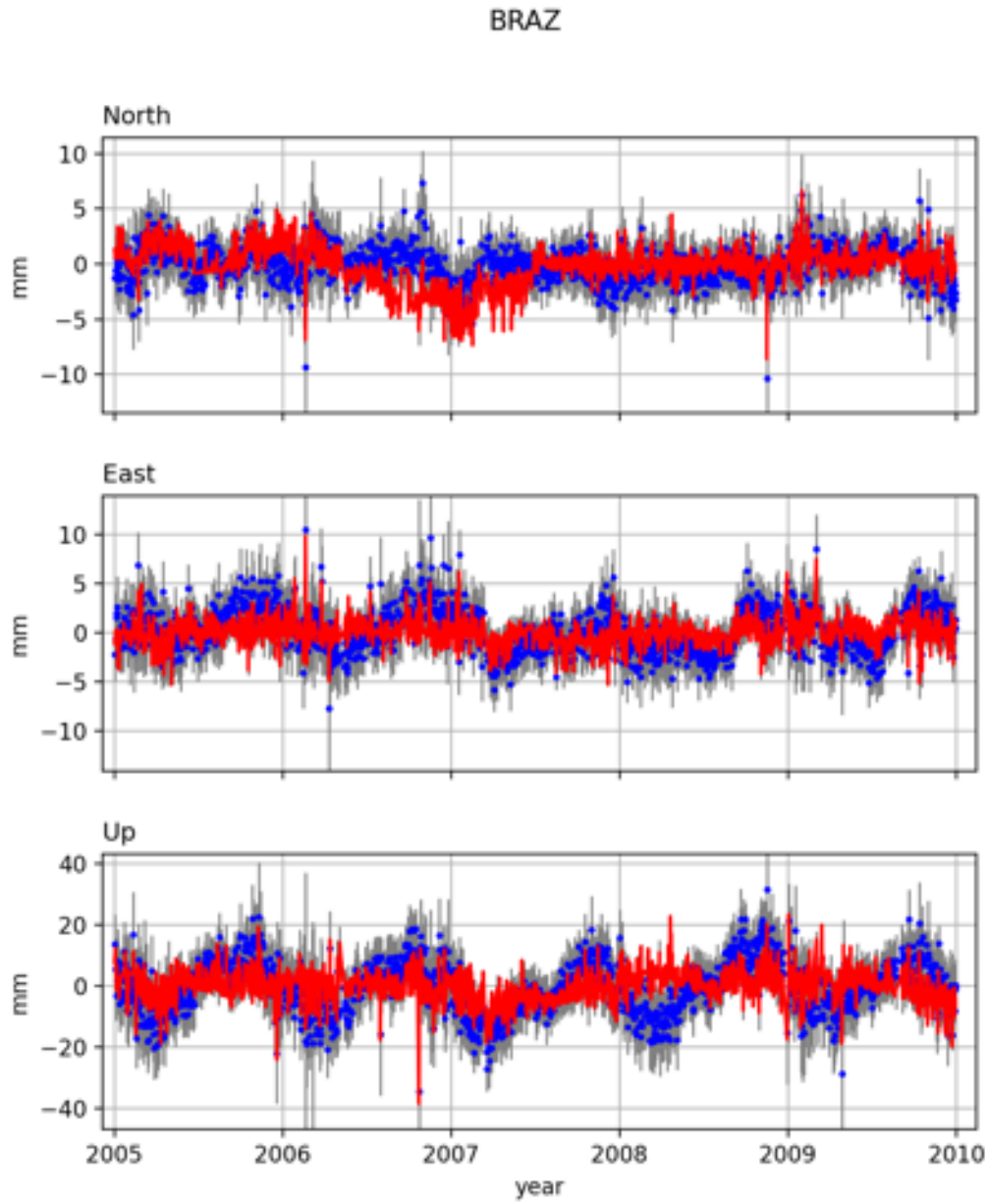
Figure 4.8: BRAZ time series (detrended) after regional common mode filtering.
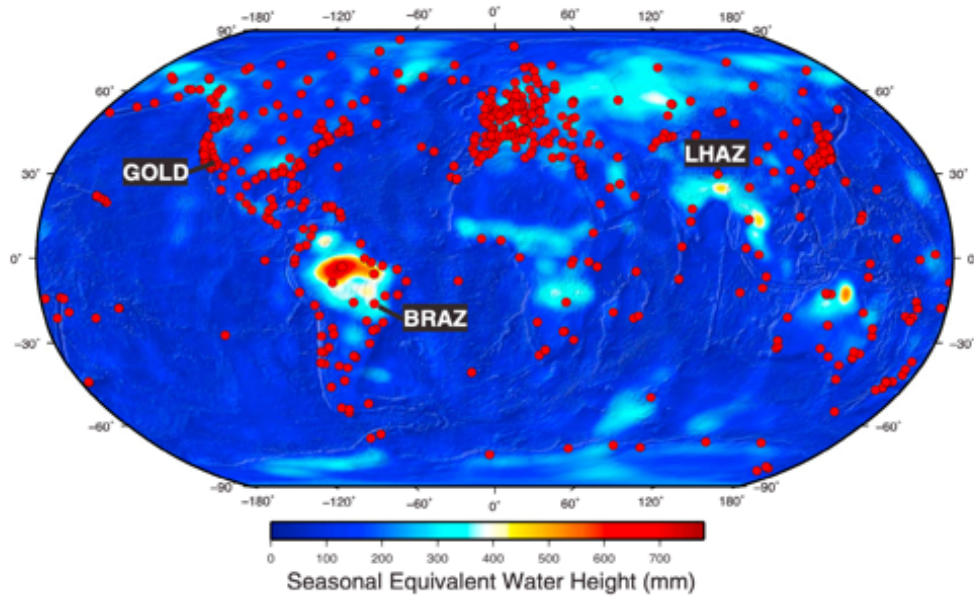
Figure 4.9: Peak-to-peak surface load variations, expressed in equivalent-water height (in mm), derived from Gravity Recovery and Climate Experiment for the 20022012 period.Red dots show the locations of the continuous IGS GPS stations. From Chanard et al. (2018)

4.6.2   Surcharge atmosphérique

Bien que lamplitude du signal annuel de la surcharge atmosphérique est généralement faible (1-2 mm, max 3 mm sur lAsie, lAntarctique et lAustralie), des études à la fois théoriques et appuyées sur des mesures indiquent que des déplacements induits par la surcharge atmosphérique peuvent atteindre 25 mm aux latitudes moyennes. De plus, il existe une énergie non-négligeable de ce signal à la période 24 h, et depuis peu, il est proposé de le prendre en compte dans le calcul du calcul journalier.
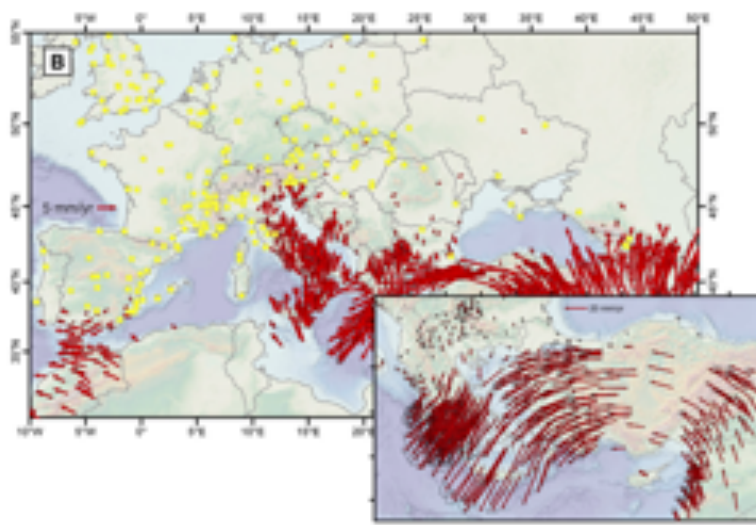
5. Geodetic Velocity fields

In this chapter, we will analyze GPS velocity fields, focusing on the horizontal components.

Dans la suite, on suppose que l'on a un champ de vitesse $v_{i \in <1; n>}$, i désignant le numéro du point, et n le nombre de points, avec C_v la matrice variance-covariance associée aux vitesses.

5.1 Equation de mouvement d'un bloc rigide sur la sphère

Si l'on considère le mouvement d'un bloc rigide sur la sphère, la vitesse horizontale d'un point M_i appartenant à ce bloc est donnée par la relation:

$$v(M_i) = \dot{\omega} \times r_{OM_i} \tag{5.1}$$

où $\dot{\omega}$ est le vecteur taux de rotation $r_{OM_i} = \begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}$ est le vecteur radial joignant l'origine de la Terre au point M_i. Trois paramètres suffisent donc à décrire la cinématique d'un bloc. L'équation (5.1) peut être utilisée de différentes manière:

1. pour déterminer la cinématique d'un bloc
2. pour prédire la vitesse relative le lonf des frontières des blocs
3. pour quantifier le niveau de déformation/rigidité d'un bloc
4. pour tenter définir de manière algorithmique la géométrie des blocs

5.1.1 Détermination de la cinématique d'un bloc

Pour chaque point M_i de vitesse $v^{i \in <1; n>}$, $v_i = \begin{bmatrix} v_{x_i} \\ v_{y_i} \\ v_{z_i} \end{bmatrix}$ appartenant au même bloc, les résultats des calculs GPS fournissent la vitesse 3D exprimée dans le référentiel global géocentrique cartésien qui est le plus souvent la réalisation courante de l'ITRF. On détermine le vecteur taux de rotation $\dot{\omega}$ tel

que

$$\sum_{i=1}^{n} \| \dot{\omega} \times r_{OM_i} - v_i \| \; soit\; minimum$$

On résoud le plus souvent cette équation par la méthode des moindres-carrés (norme L2), bien que j'utiliserai la résolution en norme L1 pour les modèles en blocs élastiques décrits plus loin. La norme minimisation de la norme L2 est immédiate, rapide et fournit une valeur directe de la matrice variance-covariance des paramètres estimés. Elle permet de plus de construire des tests statistiques d'appartenance au bloc et de détection des erreurs que je décris plus bas. Pour résoudre l'équation (5.1) par la méthode des moindres-carrés, on l'écrit sous la forme de l'équation linéaire:

$$\begin{bmatrix} 0 & z_i & -y_i \\ -z_i & 0 & x_i \\ y_i & -x_i & 0 \end{bmatrix} \begin{bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \end{bmatrix} = \begin{bmatrix} v_{x_i} \\ v_{y_i} \\ v_{z_i} \end{bmatrix} \tag{5.2}$$

$$A_i \qquad\qquad \dot{\omega} \quad = \quad v_i$$

On forme le système linéaire par assemblage

$$A\,\dot{\omega} = v \tag{5.3}$$

où $A = \begin{bmatrix} A_1 \\ : \\ A_n \end{bmatrix}$ et $V = \begin{bmatrix} v_1 \\ : \\ v_n \end{bmatrix}$.

La solution $\hat{\dot{\omega}}$ de l'équation (5.3) minimisant la quantité $R^T C_v^{-1} R$ ($R = v - A\dot{\omega}$) est

$$\hat{\dot{\omega}} = \left(A^T C_v^{-1} A \right)^{-1} A^T C_v^{-1} v$$

La matrice variance-covariance de $\hat{\dot{\omega}}$ est donnée par le théorème de Gauss:

$$C_{\hat{\dot{\omega}}} = \left(A^T C_v^{-1} A \right)^{-1} \tag{5.4}$$

L'équation (5.4) montre que la précision sur $\hat{\dot{\omega}}$ est fonction de A et C_v. La contribution de A à C_v reflète l'effet de réseau. Une ouverture plus grande du réseau (i.e. une meilleure couverture spatiale) produira des incertitudes plus petites. Bien sûr, un nombre plus grand de sites contribuera aussi à diminuer la VCV. La dépendance en C_v indique que des vitesses mieux déterminées (i.e. des termes diagonaux de C_v plus petits) amélioreront aussi la détermination de $\hat{\dot{\omega}}$.

Nous voyons aussi d'après (5.4) que l'incertitude sur $\hat{\dot{\omega}}$ ne dépend pas des résidus de l'estimation. Cette information peut être prise en compte en multipliant $C_{\hat{\dot{\omega}}}$ par le *facteur de variance a posteriori* $\hat{\sigma}_0$ (appelé reduced χ^2 le plus souvent dans les articles) défini par:

$$\hat{\sigma}_0 = \frac{R^T C_v^{-1} R}{N - p}$$

où n est le nombre d'observations et p le nombre de paramètres estimés (3 pour un modèle à un bloc). Avec cette formulation de l'équation, (5.3), $N = 3n$. Cependant, le modèle de bloc rigide ne prédit que les composantes horizontales de la vitesse (v_{e_i}, v_{n_i}). Je modifie donc l'équation (5.1) pour

définir la relation entre $\dot{\omega}$ et les composantes horizontales de la vitesse. Pour cela, on introduit la matrice rotation $T_{gen \to loc_i}$ qui transforme les coordonnées cartésiennes géocentriques dans le repère local du point M_i:

$$T_{gen \to loc_i} = \begin{bmatrix} -\sin\lambda_i & \cos\lambda_i & 0 \\ -\cos\lambda_i.\sin\phi_i & -\sin\lambda_i.\sin\phi_i & \cos\phi_i \\ \cos\lambda_i.\cos\phi_i & \sin\lambda_i.\cos\phi_i & \sin\phi_i \end{bmatrix}$$

où λ_i et ϕ_i sont respectivement la longitude et latitude du point M_i. On note H_i la sous-matrice 2x3 extraite de $T_{gen \to loc_i}$ restreintes aux deux premières lignes.

L'équation (5.2) est transformée en

$$[H_i A_i]\,\dot{\omega} = H_i v_i = \begin{bmatrix} v_{e_i} \\ v_{n_i} \end{bmatrix}$$

La matrice VCV des vitesses doit aussi être transformée dans le repère local:

$$C_v^{local} = H C_v H^T$$

où H est la matrice issue de l'assemblage des matrice individuelles H_i

$$H = \begin{bmatrix} H^1 & 0 & \cdots & 0 \\ 0 & H^2 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & H^n \end{bmatrix}$$

Cette formulation fournit le calcul correct pour $\hat{\sigma}_0$ et d'autres quantité statistiques. $\hat{\sigma}_0$ est un indicateur de la qualité de l'estimation. Si $\hat{\sigma}_0 \gg 1$, cela peut refléter soit que C_v sous-estime les incertitudes réelles des composantes de la vitesse ou que l'hypothèse de blocs rigide est incorrecte.

Le vecteur taux de rotation $\dot{\omega}$ est le plus souvent représenté sous la forme du pôle d'Euler *i.e.* par la position $(\lambda_{\dot{\omega}}, \phi_{\dot{\omega}})$ de son intersection avec la sphère terrestre et sa vitesse angulaire. Avec cette représentation, on peut représenter graphiquement la matrice VCV par une ellipse d'erreur l'ellipse d'erreur associée à $(\lambda_{\dot{\omega}}, \phi_{\dot{\omega}})$ et un écart-type sur la vitesse angulaire. Ces éléments sont données par:

$$C_{\hat{\omega}}^{local} = T_{gen \to loc} C_{\dot{\omega}} T_{gen \to loc}^T$$

Les paramètres de l'ellipse d'erreur sont:

$$\begin{aligned} \gamma &= 2\tan^{-1}(2c_{1,2}/(c_{1,1} - c_{2,2})) \\ a &= \frac{1}{2}(c_{1,1} + c_{2,2} + \sqrt{(c_{1,1} - c_{2,2})^2 + 4c_{1,2}^2} \\ b &= \frac{1}{2}(c_{1,1} + c_{2,2} - \sqrt{(c_{1,1} - c_{2,2})^2 + 4c_{1,2}^2} \\ \sigma_{\dot{\omega}} &= \sqrt{c_{3,3}} \end{aligned}$$

où γ est l'azimut du demi grand-axe, a le demi grabd-axe and b le demi petit-axe, et $\sigma_{\dot\omega}$ l'écart-type de la vitesse angulaire et les $(c_{i,j})_{(i,j)\in<1,3>^2} = C_v^{local}$. Il est important de noter que le quadruplet $(\gamma, a, b, \sigma_{\dot\omega})$ indiqués dans les articles, ne suffit pas à décrire complètement la VCV (les coefficient de corrélation $c_{3,1}$ $c_{3,2}$ ne peuvent pas être retrouvés). Ce point devient important lors de la conduction de tests statistiques.

5.1.2 Tests statistiques

Des tests statistiques peuvent être définis pour:

1. détecter des "outliers". Un outlier peut refléter soit une erreur dans les données ou de la déformation dans le bloc.
2. décider si un point présente une vitesse cohérente avec un bloc donné ou un ensemble de points
3. décider si l'addition de nouveaux blocs est nécessaire pour décrire la cinématique d'une région.

Test "F ratio" (Fisher)

Ce test a été utilisé par différents auteur pour tester les modèles cinématiques globaux Il permet à travers la comparaison de deux estimations, de décider lequel des deux modèles explique le mieux les données. Pour un ensemble d'observations, rajouter des paramètres estimés va diminuer le χ^2. Le F ratio indique si la diminution de χ^2 d'un modèle à p2 vs p1 (p1 > p2) degrés de liberté est significatif. La quantité $R^T P R$ ($P = C_v^{-1}$) suit une loi de probabilité du χ^2 à p degrés de liberté. La quantité:

$$F = \frac{[\chi^2(p1) - \chi^2(p2)]/(p1-p2)}{\chi^2(p2)/p2}$$

suit une distribution de Fisher-Snedecor à $(p1, p2)$ degrés de liberté. La valeur calculée empriquement à partir des données est donc comparée à la prédiction de la loi $F(p1 - p2, p2)$ pour un niveau de risque α donné (α =1% or 5%, correspondant à un niveau de confiance de 95% or 99%). L'hypothèse nulle dans le test est que l'ensemble des observations peut être modélisée avec un seul bloc On teste donc:

$$F \leq f_\alpha^{p1-p2,p2} \tag{5.5}$$

où $f_\alpha^{p1-p2,p2}$ est le fractile α% de la loi de Fisher-Snedecor à $(p1, p2)$ degrés de liberté. Ce test peut être utilisé de deux façons. Premièrement, il fournit un moyen de tester si un site présente une vitesse cohérente avec un autre ensemble de sites. Pour cela, on compare deux estimations, la première incluant l'ensemble des sites, la seconde, excluant le site à tester. Pour la première, on a $p1 = 2 * N - 3$ où N est le nombre total de sites. La seconde est $p2 = 2 * N - 3 - 2$ et $p1 - p2 = 2$. Si (5.5) n'est pas vérifiée, on peut en conclure avec α% de risque de se tromper que la vitesse du site testé n'est pas compatible avec la rotation estimée pour les $N - 1$ sites restants. La seconde utilisation est de tester si deux (ou plus) blocs rigides expliquent significativement mieux le champ de vitesse. Dans ce cas, deux vecteurs taux de rotation sont estimés et $p2 = 2 * N - 2 * 3$, $(p1 - p2 = 3)$. Si (5.5) n'est pas vérifiée, on peut en déduire que la diminution de χ^2 is significative et que deux blocs explique mieux le champ de vitesse. Le principal avantage de ce test par rapport à ceux décrits plus bas est qu'il est insensible à une sur- ou sous- estimation d'ensemble de la matrice variance-covariance des vitesses horizontales.

Test de Student

Le test de Student est basé sur l'analyse des résidus lors de l'estimation du vecteur taux de rotation. Il généralise le test classique du résidus normalisé en prenant en compte la variance du vecteur taux de rotation et la facteur de variance a posteriori. La quantité testée est :

$$S_i = \frac{|R_i|}{\sqrt{\hat{\sigma}_0}}\sqrt{C_{v_{ii}}} \tag{5.6}$$

où R_i est le résidu de la $i^{ème}$ observation, $\hat{\sigma}_0$ est le facteur de variance *a posteriori* et $C_{v_{ii}}$ est le $i^{ème}$ élément diagonal de la matrice variance-covariance des résidus définie par :

$$C_v = C_{v_{en}} - D(D^T C_{v_{en}} D)^{-1} D^T$$

avec $D = (R_i A_i)_{i \in <1;n>}$. S_i suit une distribution de Student à $2 * N - 3$ degrés de liberté. Le test de Student indique si S_i diffère significativement de 0. Ce test peut être conduit "in context" c'est à dire avec tous les sites inclus dans l'estimation. Il est plus discriminant en l'utilisant "out of context" c'est à dire en retirant le point considéré de l'estimation. A l'expérience, ce test fonctionne bien lorsqu'un nombre suffisant de sites ($N > 5$) est inclus dans l'estimation.

Test du χ^2

Le test du χ^2 formalise le test intuitif qu'une vitesse résiduelle est significative si elle "sort" de son ellipse d'erreur pour un niveau de confiance donné, pris usuellement de 95% ou 99%. Pour un vecteur taux de rotation $\dot{\omega}$ avec une matrice variance-covariance $C_{\dot{\omega}}$, on teste si un point M_i de vecteur vitesse horizontal exprimé dans le repère local v_{en_i} ($C_{v_{en_i}}$) présente une vitesse compatible avec celle prédite par $\dot{\omega}$. L'écart entre la vitesse horizontale prédite et celle observée est

$$v_i = v_{en_i} - H_i A_i \dot{\omega}$$

La matrive VCV associée à cette vitesse résiduelle est

$$C_{v_i} = C_{v_{en_i}} + H_i A_i C_{\dot{\omega}} (H_i A_i)^T$$

La quantité $(v_i)^T C_v^{i^{-1}} v_i$ suit une loi de distribution du χ^2 à 2 degrés de liberté. Si $(v_i)^T C_v^{i^{-1}} v_i > \varepsilon_\alpha^2$, ($\varepsilon_\alpha^2$ est le fractile d'ordre α % pour la loi du χ^2 à 2 degrés de liberté), alors la vitesse résiduelle est incluse dans son ellipse d'erreur et elle n'est pas significative.

Les tests définis précédemment permettent de définir des algorithmes de recherche de blocs semi-automatiques.

5.1.3 Evaluation du niveau de déformation interne des blocs rigides

Si l'hypothèse de bloc rigide est vérifiée, alors les résidus doivent être distribués de manière aléatoire et refléter simplement l'erreur de mesure. L'histogramme des résidus doit suivre une gaussienne. Ce test de permet cependant pas de voir une régionalisation des tendancees des résidus. Nous reviendrons sur ce point plus bas.

Un indicateur du niveau de rigidité est donné par l'écart quadratique moyen des résidus (Weighted Root Mean Square, WRMS) défini par:

$$WRMS = \sqrt{\frac{\sum_{i=1}^{n}\left(\dfrac{r_{e_i}^2}{\sigma_{e_i}^2} + \dfrac{r_{n_i}^2}{\sigma_{n_i}^2}\right)}{\sum_{i=1}^{n}\left(\dfrac{1}{\sigma_{e_i}^2} + \dfrac{1}{\sigma_{n_i}^2}\right)}}$$

où r_{e/n_i} désigne la composante e/n de la vitesse résiduelle et σ_{e/n_i} son écart-type.

5.1.4 Vitesse relative le long des frontières de blocs

Une utilisation importante des modélisations en blocs rigides est leur capacité à prédire la vitesse relative le long de leur limites. En terme d'interprétation tectonique, cette analyse fournit les conditions cinématiques aux limites d'une zone de déformation (par exemple la vitesse relative des plaques Nubie/Eurasie est la condition cinémtiques aux limmites de la Méditerranée) où la vitesse de glissement le long d'une faille majeure délimitant un bloc. Cette analyse permet de discuter quantitativement le rôle des failles dans le champ de déformation. Elle est aussi importante en terme d'évaluation de l'aléa sismique généré par un segment de faille ou une zone de déformation donnée. Si l'on suppose que l'on a deux blocs rigides 1 et 2 de vecteurs rotation $\dot{\omega}_1$ et $\dot{\omega}_2$, séparés par une faille unique idéale, composée des segments $F_j(\lambda_j, \phi_j, \lambda_{j+1}, \phi_{j+1})$ où $(\lambda_i, \phi_i, \lambda_{i+1}, \phi_{i+1})$ représentent les coordonnées géographiques des extrémités des segments. Le mouvement relatif pour un point $M_i(\lambda_i, \phi_i)$ de la faille est:

$$v(M_i) = (\dot{\omega}_1 - \dot{\omega}_2) \times r_{OM_i}$$

et si v est exprimé en coordonnées locales:

$$v(M_i) = H_i (\dot{\omega}_1 - \dot{\omega}_2) \times r_{OM_i}$$

Pour indiquer si le mouvement prédit va être dextre ou senestre, normal ou inverse, nous avons besoin de définir certaines conventions. Pour cela, on utilise les fonctions primitives implémentées dans les Systèmes d'Informations Géographiques (SIG) qui reposent sur un graphe planaire et une topologie associée. Chaque segment F_j d'une faille est composé d'un sommet initial et final, d'un domaine droite et d'un domaine gauche. L'orientation est définie en allant du sommet initial vers le sommet final. Nous pouvons définir un vecteur directeur unitaire t_j pour le segment j. n_j est le vecteur unitaire complémentaire tel que (t_j, n_j) définit un repère direct dans le plan local. Avec cette convention, les composantes parallèles et normales au segment de faille j sont:

$$
\begin{aligned}
s_j &= v(M_i).t_j \\
p_j &= v(M_i).n_j
\end{aligned}
$$

Avec cette convention, $s_j > 0$ correspond à un mouvement dextre. $p_j > 0$ correspond à un mouvement normal. Il est aussi important de calculer l'incertitude associée à ces composantes. Pour cela, on forme d'abord la matrice variance-covariance associée au vecteur taux de rotation relatif:

$$C_{1-2} = C_1 + C_2$$

ou bien si $\dot{\omega}_1$ et $\dot{\omega}_2$ ont été déterminés simultanément et présentent des termes de covariance:

$$C_{1-2} = C_1 + C_2 - C_{12} - C_{12}^T$$

où C_{12} est la matrice 3x3 incluant les termes de covariance entre les composantes de $\dot{\omega}_1$ et $\dot{\omega}_2$.

La matrice variance-covariance du vecteur taux de rotation en coordonnées locales, propagées au point M_i est:

$$C_{v_i} = H_i . A_i . C_{1-2} . \left(A_i\right)^T . \left(H_i\right)^T$$

Les incertitudes des composantes parallèles et normales sont:

$$
\begin{aligned}
\sigma_{ss_i} &= \left(t_j^T . C_{v_i} . t_j\right)^{-1} \\
\sigma_{n_i} &= \left(n_j^T . C_{v_i} . n_j\right)^{-1}
\end{aligned}
$$

5.1.5 Champ de vitesse lissé

L'identification des très faibles déformations reste un problème difficile et le plus souvent les vitesses résiduelles obtenues ne sont pas significative individuellement. On peut alors produire un champ de vitesses lissé pour voir si certaines formes du champ de vitesse sont communes à des ensemble de sites.

En effet, si les vitesses résiduelles sont distribuées de manière aléatoire, alors faire une moyenne sur une certaine zone géographique doit encore diminuer les vitesses résiduelles. Au contraire, si certaines caractéristiques du champ de vitesse sont localement cohérentes, une moyenne régionale doit permettre de mettre en évidence les tendances. Dans notre étude de la déformation du continent nord-américain, nous calculons une vitesse moyenne régionale par:

$$v = \frac{\sum_{i=1}^{N} w_i v_i}{\sum_{i=1}^{N} w_i} \tag{5.7}$$

où w_i est un poids diminuant avec la distance au point considéré.

Rotation rigide et profile de vitesse

Il est parfois utile pour illustrer la discussion tectonique de regarder les vitesses dans une direction particulière. Cependant, les variations de vitesses proviennent soit d'un effet de rotation, soit de déformation. Cette approche doit donc être utilisée avec précaution et le plus souvent sur une étendue spatiale restreinte ou l'effet de rotation peut être négligé. Un cas intéressant est celui de la composante de vitesse le long d'un grand cercle. Dans ce cas, la contribution de la rotation est constante et qu'en conséquence, toute variation de la vitesse reflète une composante de déformation. En repartant de l'équation d'un mouvement de bloc rigide sur la sphère

$$\underline{v} = \underline{\dot{\omega}} \times \underline{r} \tag{5.8}$$

on écrit

$$\underline{v} = R(\underline{\dot{\omega}} \times \hat{\underline{r}}) \tag{5.9}$$

où R est le rayon de la Terre et $\hat{\underline{r}}$ le vecteur radial unitaire. La composante de la vitesse le long du grand cercle est

$$v_n = R(\underline{\dot{\omega}} \times \hat{\underline{r}}) \cdot \hat{\underline{n}} \tag{5.10}$$

où $\underline{\hat{n}}$ est le vecteur unitaire tangent au grand cercle considéré. On considère un point fixe P appartenant au grand cercle dans la direction considérée et φ l'angle $P\hat{O}M$.

Dans ce repère, la variation de v_n en fonction de l'angle φ est:

$$\frac{\partial v_n}{\partial \varphi} = R(\underline{\dot{\omega}} \times \frac{\partial \hat{\underline{r}}}{\partial \varphi}) \cdot \underline{\hat{n}} + R(\underline{\dot{\omega}} \times \hat{\underline{r}}) \cdot \frac{\partial \underline{\hat{n}}}{\partial \varphi} \tag{5.11}$$

soit

$$\begin{aligned} \hat{\underline{r}} &= \cos\varphi\,\underline{i} + \sin\varphi\,\underline{k} \\ \underline{\hat{n}} &= -\sin\varphi\,\underline{i} + \cos\varphi\,\underline{k} \end{aligned} \tag{5.12}$$

où $\underline{i}$ et $\underline{k}$ sont les vecteurs unitaires ($\underline{i} = \frac{OP}{OP}$) et $\underline{i}.\underline{k} = 0$

$$\begin{aligned} \frac{\partial \hat{\underline{r}}}{\partial \varphi} &= -\sin\varphi\,\underline{i} + \cos\varphi\,\underline{k} = \underline{\hat{n}} \\ \frac{\partial \underline{\hat{n}}}{\partial \varphi} &= -\cos\varphi\,\underline{i} - \sin\varphi\,\underline{k} = -\hat{\underline{r}} \end{aligned} \tag{5.13}$$

$$\frac{\partial u_n}{\partial \varphi} = R(\underline{\dot{\omega}} \times \hat{\underline{n}}) \cdot \hat{\underline{n}} - R(\underline{\dot{\omega}} \times \hat{\underline{r}}) \cdot \hat{\underline{r}} = 0 \tag{5.14}$$

5.2   Blocs élastiques

5.3   Analyse en déformation

L'analyse en déformation a pour but de quantifier les variations locales du champ de vitesse. Contrairement à l'approche en blocs rigides, il ne requiert aucune hypothèse sur le type de déformation et fournit donc une alternative à l'approche en blocs rigides. Il est de plus indépendant du repère dans lequel sont exprimées les vitesses. Comme on s'interesse aux variations locales, on peut travailler en repère local, avec une approximation de Terre plate.

Définition des tenseurs gradient de vitesse et tenseur de déformation

Dans la suite, on considère un champ de vitesse horizontal $v(x, y)$ où x et y désigne les coordonnées Nord et Est. Au voisinage du point $M_0(x_0, y_0)$, on définit

$$\begin{aligned} x &= x_0 + \delta x \\ y &= y_0 + \delta y \end{aligned}$$

et $\delta X = \begin{bmatrix} \delta x \\ \delta y \end{bmatrix}$. On approxime le champ de vitesse par son développement limité au premier ordre:

$$v(x, y) \simeq v(x_0, y_0) + \nabla v(x_0, y_0). \, \delta X \tag{5.15}$$

où

$$\nabla v(x_0, y_0) = \begin{bmatrix} \frac{\partial v_x}{\partial x}(x_0, y_0) & \frac{\partial v_x}{\partial y}(x_0, y_0) \\ \frac{\partial v_y}{\partial x}(x_0, y_0) & \frac{\partial v_y}{\partial y}(x_0, y_0) \end{bmatrix}$$

∇v est le tenseur gradient de vitesse. Il est classiquement séparé en sa partie symétrique et anti-symétrique. Sa partie symétrique est le tenseur taux de déformation $\dot{\varepsilon}$ et sa partie anti-symétrique taux de rotation local $\dot{\omega}$.

$$\begin{aligned} \nabla v &= \frac{\nabla v + \nabla v^T}{2} + \frac{\nabla v - \nabla v^T}{2} \\ &= \qquad \dot{\varepsilon} \qquad + \qquad \dot{r} \end{aligned}$$

$r = \begin{bmatrix} 0 & -\dot{w} \\ \dot{w} & 0 \end{bmatrix}$ et $\dot{\omega} = \frac{1}{2}(\frac{\partial V_x}{\partial y} - \frac{\partial V_y}{\partial x})$.

$\dot{\varepsilon} = \begin{bmatrix} E_{11} & E_{12} \\ E_{12} & E_{22} \end{bmatrix}$ est symétrique et peut être diagonalisé. Les vecteurs propres correspondent aux directions de maximum de déformation. Les valeurs propres associées correspondent aux taux de déformation dans les directions des vecteurs propres:

$$\begin{aligned} \dot{\varepsilon}_1 &= \tfrac{1}{2}(E_{11} + E_{22}) + \sqrt{\Delta} \\ \dot{\varepsilon}_2 &= \tfrac{1}{2}(E_{11} + E_{22}) - \sqrt{\Delta} \end{aligned}$$

avec $\Delta = \frac{1}{4}(E_{11} - E_{22})^2 + E_{12}^2$. L'extension est avec cette formulation positive. L'azimut de la valeur propre d'extension maximum est $\varphi = -\frac{1}{2}\arctan(\frac{2E_{12}}{E_{11}-E_{22}})$. La direction perpendiculaire à φ correspond à la direction de raccourcissement maximum. La direction de cisaillement maximum est à 45° des directions propres. D'autres éléments sont interessants. $d = tr(\dot{E}) = \dot{E}_{11} + \dot{E}_{22} = \dot{\varepsilon}_1 + \dot{\varepsilon}_2 = div(v)$est la dilatation.. La dilatation est indépendante du repère et est donc le premier invariant du tenseur de déformation. Le second invariant est $\Pi = E_{12}^2 - E_{11}E_{22}$. Π joue un rôle important dans les rhéologies non-linéaires où la viscosité effective est proportionnelle à ce terme. Les taux de cisaillement $\dot{\gamma}_1, \dot{\gamma}_2$ et taux de cisaillement maximum $\dot{\gamma}$ sont :

$$\begin{aligned} \dot{\gamma}_1 &= \dot{E}_{11} - \dot{E}_{22} \\ \dot{\gamma}_2 &= 2\dot{E}_{12} \\ \dot{\gamma} &= (\dot{\gamma}_1^2 + \dot{\gamma}_1^2)^{1/2} \end{aligned}$$

.

Estimation linéaire des éléments du tenseur gradient de vitesse

On considère le champ de vitesse $v^{i=}\begin{bmatrix} v_{e_i} \\ v_{n_i} \end{bmatrix}$ avec $i \in <1; n>$ de barycentre $M_0(x_0, y_0)$. Si l'on fait l'hypothèse que le tenseur gradient de vitesse est constant sur la zone considérée, on peut estimer ses éléments par l'équation (5.15):

$$\underbrace{\begin{bmatrix} 1 & 0 & x_i - x^0 & y_i - y^0 & 0 & 0 \\ 0 & 1 & 0 & 0 & x_i - x^0 & y_i - y^0 \end{bmatrix}}_{A_i} \underbrace{\begin{bmatrix} v_x^0 \\ v_y^0 \\ \frac{\partial v_x}{\partial x} \\ \frac{\partial v_x}{\partial y} \\ \frac{\partial v_y}{\partial x} \\ \frac{\partial v_y}{\partial y} \end{bmatrix}}_{V_i} = \underbrace{\begin{bmatrix} v_{x_i} \\ v_{y_i} \end{bmatrix}}_{v_i} \qquad (5.16)$$

qui peut être résolu par moindres-carrés. Comme 6 paramètres doivent être estimés, un minimum de 3 vitesses horizontales (6 observations) est nécessaire au calcul des éléments. Le calcul le plus simple est donc obtenu avec une partition de l'espace en triangles. Cette partition est non-unique, mais une triangulation de Delauney est un choix pratique. Cette partition possède la propriété que tout cercle passant par les 3 sommets du triangles ne contient pas d'autres points. Les triangles sont les plus équilatéraux possibles.
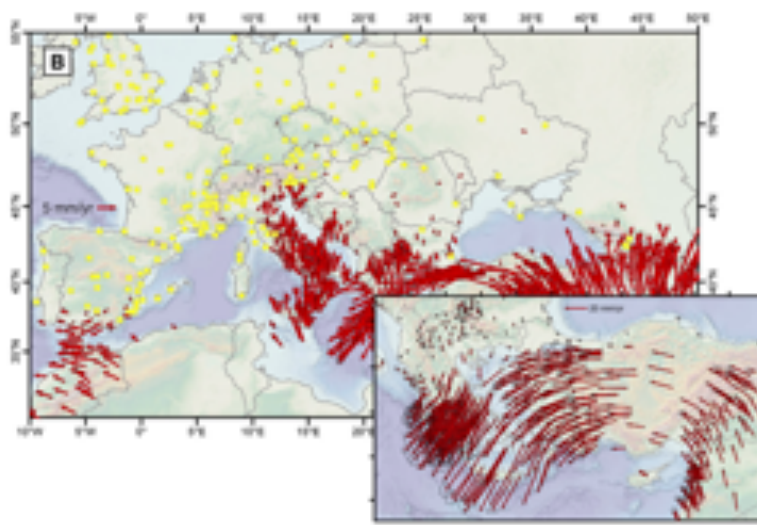
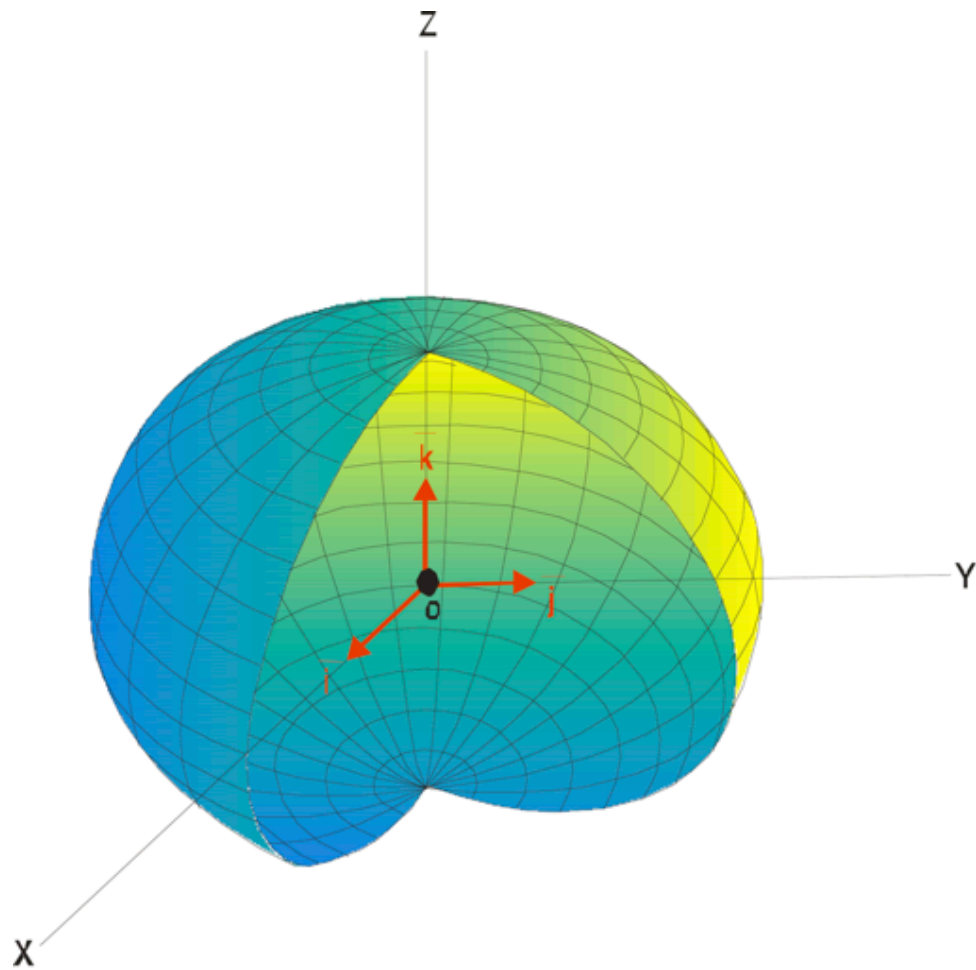# 6. Appendices

6.1   Shape of the Earth and positioning concepts

Figure 6.1: The revolution ellipsoid and geocentric frame