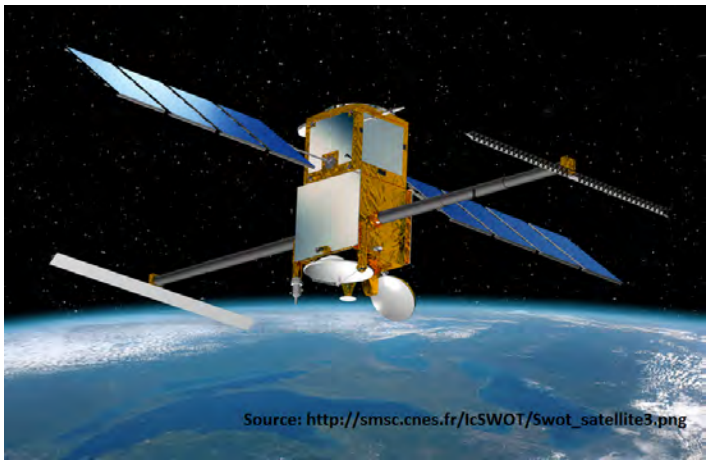
University Stuttgart
Geodetic Institute

Literature Analysis of SWOT Mission from geodetic perspective



Term Paper in Course of Studies
Geodesy and Geoinformatics
in the University Stuttgart

Wenxi Cao

Stuttgart, May 2012

Adviser: Prof. Dr.-Ing. Nico Sneeuw
University Stuttgart

Erklärung der Urheberschaft

Ich erklare hiermit an Eides statt, dass ich die vorliegende Arbeit ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt ubernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ahnlicher Form in keiner anderen Prüfungsbehorde vorgelegt und auch noch nicht veroffentlicht.

Ort, Datum Unterschrift

Satellite radar altimeter has been used for nearly twenty years to observe the variety of the global ocean surface topography. It has advanced our understanding of global ocean circulation and sea level change. However the conventional radar altimeter can not resolve the submesoscale features in the oceans because of its large spacing between satellite ground tracks and coarse ground resolution. On the other hand altimetry technique is expected to be able to observe large rivers, lakes and monitor the storage of freshwater on land. These new challenges require a new technique and a new mission. In 2016 a satellite mission called **S**urface **W**ater and **O**cean **T**opography (SWOT) will be launched by NASA and CNES according to plan. This term paper will summarize the general measurement principle, orbit design issues and applications of SWOT in the literature.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Geodesy is the scientific discipline that deals with the measurement and representation of the Earth, including its gravitational field, in a three-dimensional time-varying space. In order to acquire geometrical information on the Earth's surface, different series of sensors and platforms will be used to perform observation according to applications. One type of them is called satellite radar altimeter, which has been used for nearly twenty years to observe the variety of the global ocean surface topography. Occasionally it has been also used to help the hydrologic study on lands. This instrument can transmit radio waves towards the ground in the nadir direction of satellite and record the time elapse of the echoes. Along with the wave analysis the altitude of the satellite can be accurately calculated (*Wikipedia, the free encyclopedia*, n.d.).

Until now precise altimetry missions like TOPEX/Poseidon, Jason-1, and OSTM / Jason-2 and other scientific multi-mission projects have provided observations of the ocean surface topography at scales larger than about 200 km. These observations made significant advances in our understanding of global ocean circulation and sea level change. These large-scale measurements also provided us with information about heat storage of the global oceans and processes underlying the changes of the ocean in relation to climate cycles such as El Nino and La Nina. With these observations ocean models can be approved from coarse-resolution to higher resolutions, we are now able to produce simulations of the present state of the ocean which compare increasingly well to observations (Fu et al., 2010a).

Although the success of satellite radar altimetry in the scientific study of oceanography at mesoscale, this technique encounters challenges as the study of oceanography goes deep into the submesoscales (scale $<$ 100 km), which refer to 90% kinetic energy of the oceans, ocean variability in the form of fronts and filaments and global heat and CO_2 circulation even biochemical cycle of the ocean. Despite of this importance of submesoscale features of the ocean, the conventional radar altimeter can not satisfy the requirements for the scientific study in the oceanography because of its large spacing between satellite ground tracks and coarse ground resolution (about 200 km). On the other hand altimetry technique is more and more desired to apply to hydrology. It is expected to be able to observe large rivers, lakes and monitor the storage of freshwater on land (Fu et al., 2010a).

In order to fulfill there oceanographic and hydrologic requirements a new satellite mission called **S**urface **W**ater and **O**cean **T**opography (SWOT) is being developed for surface water observation in the future. The SWOT mission uses a different measurement principle. It uses InSAR technique which needs two antennae to measure not only altitude of points, but also phase information of the echoes. For this reason a Ka-band radar interferometer is loaded on the satellite to make high-resolution wide- swath altimetry measurements. The SWOT mission

is a collaborative project between NASA and CNES, it is anticipated to be launched in 2016 (Fu et al., 2010a).

In this term paper a general literature research will be given, including explanation of the measurement principle, orbit design, function of radiometer and some applications of SWOT mission. These contents will be divided into four chapters:

1. The first chapter is devoted to explanation of InSAR principle. A general review of radar technique is provided so that we can see how developed InSAR from the conventional radar and how is altitude information calculated from measurements by InSAR. At the end of this chapter an error analysis will be given.
2. The second chapter deals with orbit design issues. We firstly review the orbit design for other altimetry missions such as TOPEX to obtain a general approach. We will see that the orbit design for SWOT is based on the similar approach. A final orbit design from literature is given at the end of this chapter. In this chapter a Matlab program for orbit simulation is written to verify the orbit parameters found in the literature.
3. In the third chapter the focus is placed on the radiometer for SWOT. The measurement principle of the radiometer for SWOT is similar to that in the previous missions. They are both statistical algorithms. An analytical derivation is also provided to explain the physical background of statistical algorithms.
4. The last chapter is for the applications of SWOT in oceanography and hydrology.
5. An appendix is at the end including a derivation of the radiative transfer equation used in the third chapter and Matlab codes used in the second chapter.

Chapter 2

Ka-band interferometric radar

2.1 Introduction

Ka-band **IN**terferometric **R**adar (KaRIN) is the main measurement instrument of the **S**urface **W**ater **O**cean **T**opography (SWOT) mission. KaRIN consists of following instruments (Fu et al., 2010a):

- an interferometer, which uses InSAR technique to produce high accuracy height measurements,
- a microwave radiometer similar to the advanced microwave radiometer on Topex/Poseidon mission, used to provide range corrections due to wet atmospheric delay,
- a nadir looking conventional altimeter used to fulfill the gap between swaths and calibrate the interferometer.
- In addition, a **P**recice **O**rbit **D**etermination system (POD) on KaRIN can ensure the accuracy of orbit positioning on the order of cm. POD consists of three basic positioning system: **G**lobal **P**ositioning **S**ystem (GPS), **D**oppler **O**rbitography and **R**adiopositioning **I**ntegrated by **S**atellite (DORIS), and laser ranging system.

KaRIN is an updated version of **In**terferometric **S**ynthetic **A**perture **R**adar (InSAR) technique which has found many applications in the last decade. It is characterized by its ability to produce high resolution wide swaths (Fu et al., 2010a). The traditional InSAR uses complex signals measured by two antennas separated by a known distance in space, one antenna radiates energy to the Earth's surface and both of them receive the echoes. From the phase comparison and triangulation the height of target points on the Earth's surface is determinable.

SWOT inherits the InSAR principle of previous missions such as **S**huttle **R**adar **T**opography **M**ission (SRTM) and **W**ide **S**wath **O**cean **A**ltimeter (WSOA) (Fu et al., 2010a). In comparison to these space missions, some instrumental parameters have to be changed by SWOT to fulfill its scientific work. The look angle is limited to about 4.5° to reduce the measurement error due to error in baseline roll angle. As a result of reduction of look angle, the ground swath is also reduced, so two ground swaths are designed to compensate the effect of the small look angle. In order to distinguish these two ground swaths, beams have to be differently polarized.

In this chapter the InSAR principle of KaRIN in SWOT will be briefly discussed, some procedures of data processing such as co-registration, spectral shift compensation, phase unwrapping and discussion of ambiguity issues are neglected. We are only concentrated on the main measuring principle of InSAR and error budget of instrument in the mission design. The part

associated with InSAR has referenced the work (Soumekh, 1999; Rosen et al., 2000; Hassen, 2001; Ferretti et al., 2007). The part associated with error budget has referenced the work (Rodriguez et al., 1999).

2.2 Radar resolution

Radar resolution describes the ability of radar to observe details on the Earth's surface. It is labeled by radial or ground resolution, and the resolution in the azimuth dimension.

2.2.1 Radial dimension

According to the definition by Skolnik (1970), **radar** is an abbreviation for **Ra**dio **D**etection **a**nd **R**anging, which is an electromagnetic sensor for the detection and location of reflecting objects. Radar can radiate electromagnetic energy from an antenna in space, a part of this energy will then be reflected by a target far away from the radar and again received by the radar. Because of attenuation of energy due to medium in space, the reflected energy is tiny in comparison with the radiating energy. The received energy will then be amplified by a receiver. As the electromagnetic energy propagates in space, there are always clutters and noises disturbing the useful signal, the received signal can be mixed with these clutters and noises, an adequate signal processing is necessary to detect and identify the target location. In a simplified way the whole process can be illustrated as follows.

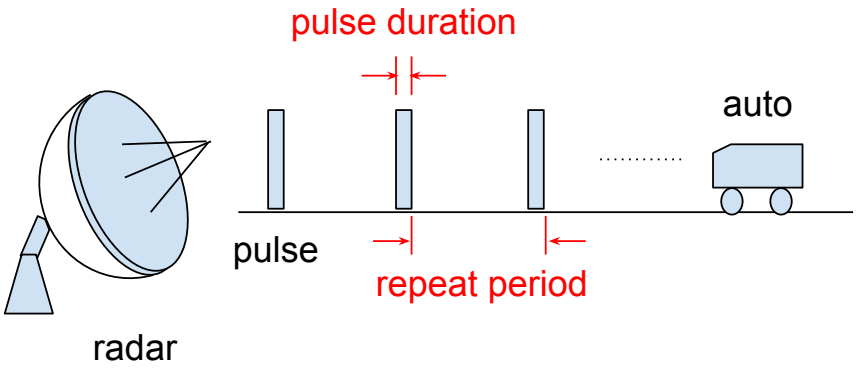


Figure 2.1: *Pulse duration and repeat period*

In the figure 2.1, radar transmits a rectangular pulse with duration time τ periodically. The repeat period is T_p which is always called *pulse repeat interval (PRI)*, an important parameter

in the radar design is defined as $1/T_p$ or $1/\mathrm{PRI}$, called *pulse repeat frequency (PRF)*, it plays an important role in ambiguity issue.

As the electromagnetic wave propagates in space, it will be reflected when it touches an object. If the time elapse between transmission and reception of the electromagnetic wave is denoted as t, so the *distance* or *range* from the radar to the target is:

$$\rho = \frac{1}{2}ct. \tag{2.1}$$

But the objects within the pulse duration can not be distinguished from each other, so the *radial resolution* of radar can be defined as:

$$\delta\rho = \frac{1}{2}c\tau. \tag{2.2}$$

2.2.2 Resolution on ground

Radial dimension is the dimension where the signal propagates and it is also called *slant range dimension*. The space-borne radar is always side-looking in order to avoid ambiguity, so that the radial resolution will be projected on the ground and not coincident with the ground resolution. This situation can be seen in the figure 2.2.

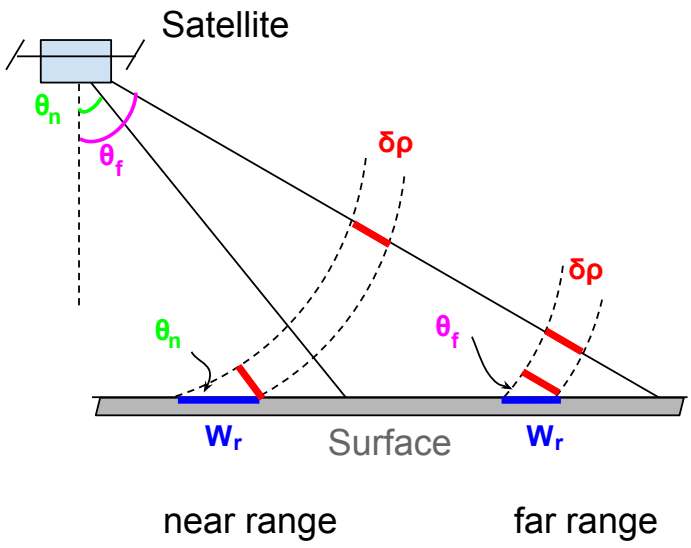


Figure 2.2: *Ground resolution*

In the figure 2.2 we see that the pulse lasts only a short period $\delta\rho$. As the pulse transmits in far range direction or in near range direction, the look angel θ can be different, which determines the ground resolution W_r. If we assume that the satellite is high enough above the Earth's surface, then the pulse can be considered as parallel transmitted, so we have a simple geometrical relationship between ground resolution and look angel:

$$W_r = \frac{\delta\rho}{\sin\theta}. \tag{2.3}$$

From this formula we see that the ground resolution is smaller at the far range than that at the near range.

2.2.3 Azimuth dimension

As discussed in previous subsections, radar can be used to measure distance of a target. On the other hand, the electromagnetic beam of radar is always adjusted towards the target in a given direction. This configuration enables radar to perform measurements to identify targets in azimuth dimension. The basic concept is the antenna pattern.

A rectangular antenna with length L and width D and uniform current density produces a *antenna pattern* as follows:

$$\operatorname{sinc}^2\left(\frac{D}{\lambda}\phi_{\mathrm{r}}\right)\operatorname{sinc}^2\left(\frac{L}{\lambda}\phi_{\mathrm{a}}\right). \tag{2.4}$$

This formula describes the radiated energy density as a function of off center beam angle ϕ. Here ϕ_{r}, ϕ_{a} mean the beam angles in radial and azimuth direction respectively. If they are both zero, i.e. in the center of beam direction, the energy density is maximal. This direction is called *bore-sight*.

From this mathematical relationship we see that the energy density has more than one zero value points, which results in many lobes in different radiation direction. The lobe which includes the bore-sight direction is called *main lobe*, the other lobes are called *side lobes*. These side lobe can produce disturbing signals and they are always suppressed by a weighting *window function* in the antenna pattern, so in a simplified view only the main lobe is responsible for the measurement.

In the azimuth direction, the targets within the main lobe will reflect the electromagnetic energy radiated by radar, whereas the objects outside the main lobe will not, in this way targets in the azimuth dimension can be identified by radar. But echoes from the targets which are at the same distance and within the main lobe are received by the antenna at the same time, so they can not be distinguished, consequently the *azimuth resolution* is defined as the width of the main lobe whose edge is described as 3 dB decrease of maximal energy in the antenna pattern:

$$\beta_{\mathrm{a}} = 0.886\frac{\lambda}{L} \approx \frac{\lambda}{L}, \tag{2.5}$$

when this azimuth resolution is translated into a metric scale, we have:

$$W_{\mathrm{a}} = \frac{\lambda}{L}\rho. \tag{2.6}$$

2.3 Improvement of radar resolution by SAR principle

As described in the section (2.2), the radar resolution is limited by pulse duration τ in radial dimension (see equation (2.2), (2.3)), and by antenna length L in the azimuth dimension (see equation (2.6)). It is very hard to improve the radar resolutions by changing τ, L directly because of technical difficulties and limitations. These limitations can be overcome through a signal processing technique, which is known as **S**ynthetic **A**perture **R**adar (SAR) technique. The improvement is carried out in two dimensions respectively.

2.3.1 Improvement of resolution in radial dimension

As discussed in the subsection 2.2.1, the radial dimension depends on the pulse duration τ. In order to acquire a sharp resolution the pulse duration must be as small as possible. But on the other hand decrease of pulse duration means the decrease of radiated energy, if the radiative energy is too small the echo from targets can then not be extracted from background noise. So the radial resolution is limited in this way.

A transformation of formula (2.2) provides a way out this problem:

$$\delta\rho = \frac{1}{2}c\tau \approx \frac{1}{2}\frac{c}{B}, \tag{2.7}$$

where $B \approx 1/\tau$ is called *bandwidth* of the transmitted wave. This formula means that the radial resolution can be improved by increase of bandwidth of the transmitted wave. A commonly used waveform is a linear frequency modulation called *chirp* signal in a rectangular wave. This chirp signal will complicate the echoes from targets so an adequate processing should be adopted.

If the *target signal* is ideally modeled as $f(t) = \sum_n \sigma_n \delta(t - t_n)$, where σ_n is the *reflectivity* of the n-th target which is located at the position $x_n = ct_n/2$, the radiated wave is denoted as $p(t)$, the *echo signal* can then be denoted as a convolution of both:

$$s(t) = f(t) * p(t), \tag{2.8}$$

so the acquired processing can be seen as a deconvolution of target signal $f(t)$ from $s(t)$. This operation is easily performed in the frequency domain. The convolution above can be equivalently transformed as:

$$S(\omega) = F(\omega)P(\omega). \tag{2.9}$$

If the $P(\omega)$ is nonzero for all ω, $F(\omega)$ can then easily be calculated by a division operation followed by calculation of $f(t)$. But the condition that $P(\omega)$ is nonzero for all ω means an infinite bandwidth of electromagnetic wave. This condition can not be fulfilled in reality.

Another method to solve this problem is called *matched filtering*. Deriving from Equation (2.9), we have:

$$S_{\mathrm{M}}(\omega) = S(\omega)P^*(\omega) = F(\omega)P(\omega)P^*(\omega) = F(\omega)|P(\omega)|^2, \tag{2.10}$$

where $S_{\mathrm{M}}(\omega)$ is the matched processed signal, $s_{\mathrm{M}}(t)$ is its counterpart in the time domain. If the inverse Fourier transformation of $|P(\omega)|^2$ is denoted as $\mathrm{psf}(t)$, then the equation above can be rewritten as:

$$s_{\mathrm{M}}(t) = f(t) * \mathrm{pdf}(t) = \sum_n \sigma_n \delta(t - t_n) * \mathrm{psf}(t) = \sum_n \sigma_n \mathrm{psf}(t - t_n). \tag{2.11}$$

At the end, the ideal modeled target signal

$$f(t) = \sum_n \sigma_n \delta(t - t_n) \tag{2.12}$$

is received and identified as

$$s_{\mathrm{M}}(t) = \sum_n \sigma_n \mathrm{psf}(t - t_n). \tag{2.13}$$

The difference between $f(t)$ and $s_M(t)$ is difference between the function $\delta(t - t_n)$ and function psf(t). The latter one is called *point spread function* and is also a pulse with a narrow width which depends on the frequency bandwidth of radiated electromagnetic wave. This means the matched processed signal $s_M(t)$ is more similar to the target signal $f(t)$ if the frequency bandwidth is larger.

An other viewpoint of equation (2.10) is in time domain:

$$s_M(t) = s(t) * p^*(-t) = \int_{-\infty}^{\infty} s(u)p^*(u - t)du. \tag{2.14}$$

The last form of $s_M(t)$ is a *complex correlation coefficient*. From this viewpoint we see that the matched filtering is equivalent to the calculation of a (complex) correlation coefficient. This situation is similar to the PRN code in GPS application, the time elapse of signals from GPS satellite to receiver is also calculated by correlation technique. The advantage of matched filtering is the optimal SNR (signal noise ratio), a proof can be found in the literature (Curlander and McDonough, 1991).

2.3.2 Improvement of azimuth resolution

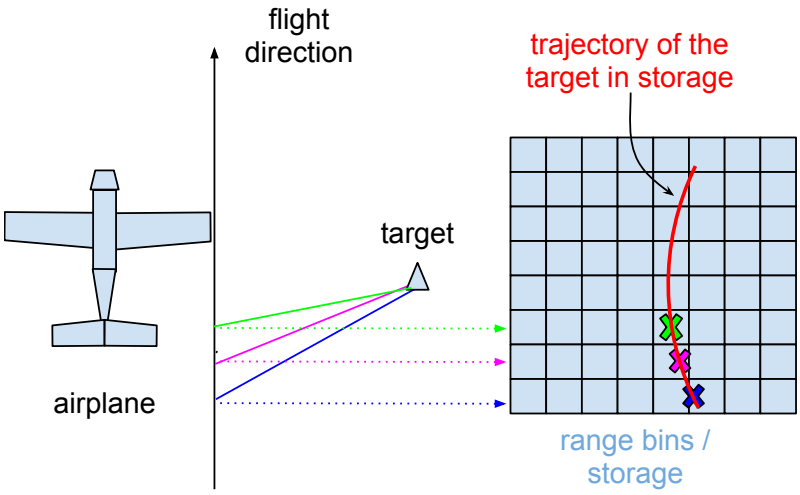


Figure 2.3: *Range migration*

As discussed in the subsection 2.2.3, the azimuth resolution depends on the distance, wavelength and length of antenna. The problem of real aperture radar is that the azimuth resolution is proportional to the distance, so the azimuth resolution is hard to limit under a given value if the radar operates on a space-borne platform. The discussion in the subsection (2.3.1) suggests that the increase of bandwidth can improve the radial resolution. The same idea can also be used to improve the azimuth resolution. In the azimuth direction, the movement of radar can arouse the Doppler effect. As the radar approaches the target, the Doppler frequency

will increase, as the radar moves away from the target, the Doppler frequency will decrease. This phenomenon produces a frequency modulated electromagnetic wave in azimuth dimension, whereas in radial dimension a frequency modulated wave has to be generated artificially. However, the idea is same: increase the bandwidth of the radiated signal.

Similar to the matched filtering in radial dimension, there must be also a processing to extract the target signal from echoes. The mathematical description of how this works is too difficult and beyond the scope of this work. The main problem that makes it more difficult is called *range migration*. As the platform moves above a target, the received echoes will be saved in an internal storage called *range bins*. This is a two dimensional (it can also be multi-dimensional) storage to save the echoes, as illustrated in the figure 2.3.

The vertical dimension of range bins is to save echoes at different platform positions. The horizontal dimension of range bins is to save echoes at the same platform position but for different targets. As the platform moves, the same target will be detected and its echo will be saved at different storage unit in the range bins, their positions in storage are illustrated as crosses in the figure. Because the distance from the platform to the target becomes smaller at the beginning but at the end becomes larger, the trajectory of crosses is similar to a parabola curve. Consequently the echo signal in azimuth dimension is coupled with the radial dimension. In order to extract azimuth echoes from range bins, we should decouple azimuth dimension from radial dimension. There are a few SAR processing algorithms in the literature, such as range Doppler algorithm, chirp scaling algorithm, Omega-K algorithm, etc. A detailed description can be found in the literature (Cumming and Wong, 2005).

2.4 SAR mode

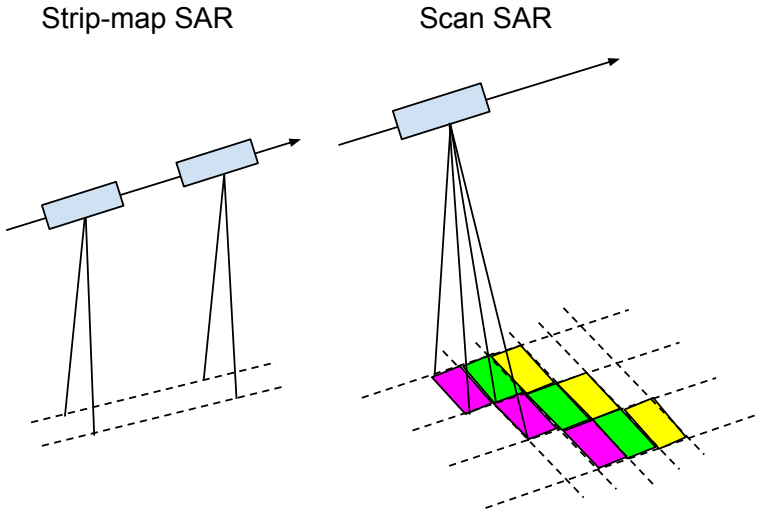


Figure 2.4: *SAR operation mode*

In the literature, there are two SAR modes, one is called *strip-map SAR*, another is called *scan SAR*. The first one is commonly used in the satellite imaging system, in this mode the beam from the radar is directed in a given direction towards Earth's surface, which can be scanned as the platform (satellite) moves. In the second mode, the beam is not directed in a constant direction but steerable. The beam can move cross the flight direction periodically, in this way a larger synthetic swath can be generated which results in the improvement of coverage. A illustration is given in the figure 2.4. The scan SAR mode in this figure has three subswaths.

In the literature (Esteban-Fernandez et al., n.d.) a prototype of onboard SAR processor is designed for SWOT. According to the description by them, this processor performance a *Doppler-sharpened, multi-squint SAR processing*, they say this approach is similar to beam forming technique, where a number of beams are created within the real aperture azimuth antenna pattern. A detailed description is at the moment not found in the literature.

2.5 Interferometric SAR

One radar can only perform two dimensional measurements. If three dimensional measurements are desired, for example positioning a point on the Earth's surface, two antenna should be combined together to accomplish the work. A commonly used technique is called *interferometric SAR*, which combines the measurements by two separated antennas with a known distance, the position of target on the Earth's surface can then be determined through triangulation calculation. There are two ways to formulate this technique. The first one solves the problem directly from distance measurements, whereas the second one solves the problem indirectly using phase information.

2.5.1 Geometrical solution

The figure (2.5) is the configuration of two separated satellites S_1 and S_2 in space which observe the same point on the Earth's surface. The distance between two satellites is B, this parameter is also called *baseline*. The direction vector formed by S_1 and S_2 deviates from the horizontal by a *roll angle* α. The angle θ is called *look angle*. The distances from satellites to the point P are the measurement values and denoted as ρ_1 and ρ_2 respectively. Here baseline length B, positions of satellites S_1 and S_2 are known parameters, measurements are distances ρ_1, ρ_2, position of point P is unknown.

In the triangle which consists of points: P, S_1 and S_2, we have following equation from the cosine rule:

$$\rho_2^2 = B^2 + \rho_1^2 - 2B\rho_1 \cos(90^\circ - \theta + \alpha), \tag{2.15}$$

then the look angle θ is resolvable:

$$\theta = 90^\circ + \alpha - \arccos\left(\frac{B^2 + \rho_1^2 - \rho_2^2}{2B\rho_1}\right). \tag{2.16}$$

Because the position of satellite S_1 and the distance between S_1 and P are known, the direction of vector from S_1 to P is also calculable from θ and α, then the position of point P can be calculated by polar point determination:

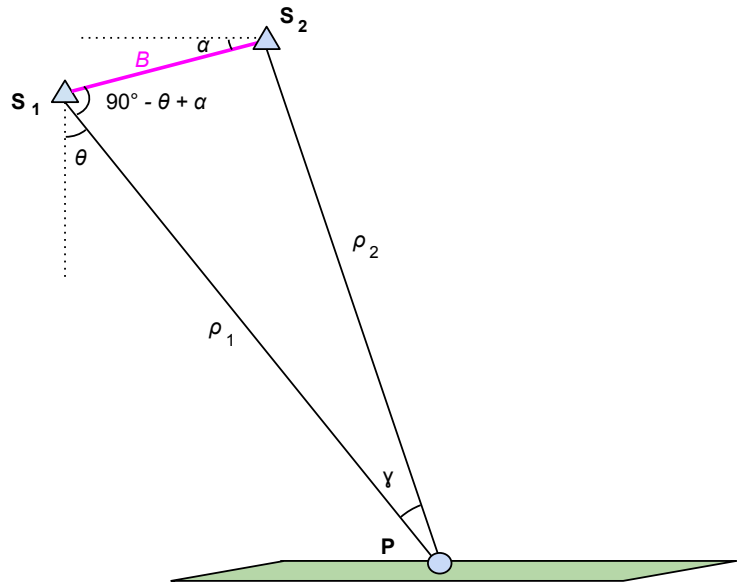


Figure 2.5: *Satellite configuration for geometrical solution*

$$\begin{aligned} x_{\mathrm{P}} &= x_{\mathrm{S}_1} + \rho_1 \cdot \sin\theta, \\ y_{\mathrm{P}} &= y_{\mathrm{S}_1}, \\ z_{\mathrm{P}} &= z_{\mathrm{S}_1} - \rho_1 \cdot \cos\theta. \end{aligned} \tag{2.17}$$

In this situation we derive the position of the point P from two known points S_1, S_2 and two distance measurements ρ_1, ρ_2, this type of problem is similar to Resection (German: "Bogenschnitt") in geodesy. A usual method of describing the quality of resection is called *Helmert point error* (German: "Helmert'scher Punktfehler"). Assuming that measurement ρ_1 is as accurate as measurement ρ_2 and both of them are uncorrelated, the Helmert point error has following form:

$$\sigma_{\mathrm{P}} = \sqrt{2}\sigma_\rho \frac{1}{\sin\gamma}, \tag{2.18}$$

where σ_{P} is the Helmert point error, σ_ρ is the distance measurement error, γ is the intersection angle of two direction vectors. Based on the fact that the altitude of SWOT satellite is about 970 km and the baseline length is about 10 m, the intersection angle will be only 10^{-5} rad, which means that the point error will be about 100 m even if the distance measurement is 1 mm accurate.

We can also estimate how large is the height error due to the Helmert point error. According to the equations (2.17) the height error can be calculated if we differentiate both sides of the third equation, then we have:

$$\sigma_{z_{\mathrm{p}}} = \sigma_{z_{\mathrm{S}_1}} - \sigma_{\rho_1} \cdot \cos\theta + \rho_1 \cdot \sin\theta \cdot \sigma_\theta. \tag{2.19}$$

We can ignore errors $\sigma_{z_{\mathrm{S}_1}}, \sigma_{\rho_1}$, only see how large is the height error from the error in look angle, so the equation above is simplified as:

$$\sigma_{z_{\mathrm{p}}} = \rho_1 \cdot \sin\theta \cdot \sigma_\theta = \Delta x_1 \cdot \sigma_\theta. \tag{2.20}$$

At the other hand the error in look angle due to the Helmert point error can be illustrated in the figure 2.6:

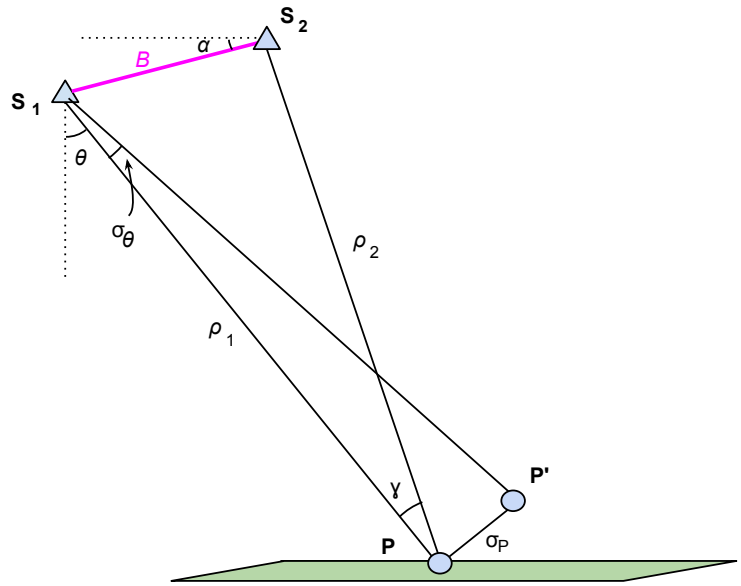


Figure 2.6: *Error in look angle due to Helmert point error*

We see that the point P on the Earth's surface is displaced to the point P', so the displacement results in an error in look angle:

$$\sigma_\theta \approx \frac{\sigma_\mathrm{P}}{\rho_1}. \tag{2.21}$$

In order to simplify the derivation, $\sin\gamma$ in the equation (2.18) can be rewritten as:

$$\sin\gamma \approx \frac{B}{\rho_1}. \tag{2.22}$$

From equations (2.18), (2.20), (2.21) and (2.22) we have:

$$\sigma_{z_\mathrm{p}} = \Delta x_1 \cdot \sigma_\theta = \Delta x_1 \cdot \frac{\sigma_\mathrm{P}}{\rho_1} \tag{2.23}$$

$$= \Delta x_1 \cdot \frac{\sqrt{2}\sigma_\rho}{\rho_1 \sin\gamma} = \Delta x_1 \cdot \frac{\sqrt{2}\sigma_\rho}{\rho_1 \frac{B}{\rho_1}} \tag{2.24}$$

$$= \sqrt{2} \cdot \frac{\Delta x_1}{B} \cdot \sigma_\rho. \tag{2.25}$$

For SWOT mission, baseline length B is about 10 m, Δx_1 can reach to half length of the swath maximally, it is about 65 km. So the height error is about 9 m even if the distance is 1 mm accurate measured. This is far from the tolerance of the mission.

2.5.2 Phase solution

The drawback of the method above is that the distance measurements are not accurate enough to perform three dimensional measurement because the baseline is rather smaller than the distance between satellite and Earth's surface. We can also see that from equation (2.25). Baseline B should be orders of magnitude larger to reduce the total error. But it is no easy to expand the baseline to this extent. A solution for it is to use the phase information recorded by the antenna.

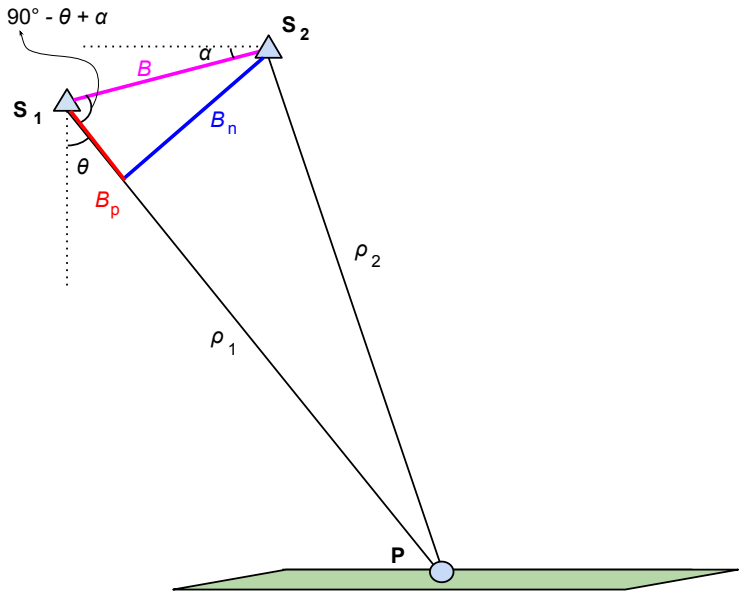


Figure 2.7: *Satellite configuration for phase solution*

Interferogram

Both of amplitude and phase information of the echo from an arbitrary point are simultaneously recorded by two antenna, denoted as $u_1(x,y)$ from antenna 1, $u_2(x,y)$ from antenna 2, they are both complex signals. The phase is related to the distance by the following formula:

$$\varphi = \frac{4\pi}{\lambda}\rho. \tag{2.26}$$

Because the pulse is transmitted towards Earth's surface and again reflected back, the pulse cover a distance twice as long as the geometrical distance, so the factor which transforms the distance to phase in the equation above is $4\pi/\lambda$ instead of $2\pi/\lambda$.

The phase difference between two complex signals is called *interferometric phase*. It is commonly extracted from *interferogram*, which is calculated by conjugation operation of two complex signals:

$$u(x,y) = u_1(x,y)u_1^*(x,y) = |u(x,y)|\exp(j\phi), \tag{2.27}$$

so the interferometric phase ϕ is included in interferogram $u(x, y)$.

This phase difference (interferometric phase) is calculable and can be considered as a known parameter. It can be used through a triangulation relationship to calculate the look angle of satellite S_1, in this way the three dimensional position of the point on the Earth's surface can be determined.

Inserting (2.26) into (2.27) gives us a relation between the interferometric phase and geometrical distance:

$$\phi = \varphi_1 - \varphi_2 = \frac{4\pi}{\lambda}(\rho_1 - \rho_2) = \frac{4\pi}{\lambda}\Delta\rho. \tag{2.28}$$

$\Delta\rho$ will be related to the look angle in the following part.

Far-field approximation

Similar to the equation (2.15), we have following relationship from the cosine rule:

$$\cos(90^\circ - \theta + \alpha) = \frac{B^2 + \rho_1^2 - \rho_2^2}{2B\rho_1}. \tag{2.29}$$

ρ_1 can also be rewritten as:

$$\rho_1 = \rho_2 + \Delta\rho. \tag{2.30}$$

Inserting the equation (2.30) into (2.29) gives:

$$\cos(90^\circ - \theta + \alpha) = \frac{B^2 + \rho_1^2 - \rho_2^2}{2B\rho_1} = \frac{B^2 + (\rho_2 + \Delta\rho)^2 - \rho_2^2}{2B\rho_1} \tag{2.31}$$

$$= \frac{B^2 + \Delta\rho^2 + 2\rho_2\Delta\rho}{2B\rho_1}. \tag{2.32}$$

Because $B, \Delta\rho$ are much smaller than ρ_1, they can be neglected in the equation (2.32):

$$\cos(90^\circ - \theta + \alpha) = \frac{B^2 + \Delta\rho^2 + 2\rho_2\Delta\rho}{2B\rho_1} \approx \frac{2\rho_2}{2\rho_1} \cdot \frac{\Delta\rho}{B}. \tag{2.33}$$

The distances ρ_1, ρ_2 are nearly same, so we have:

$$\cos(90^\circ - \theta + \alpha) \approx \frac{\Delta\rho}{B}. \tag{2.34}$$

From the configuration in the figure 2.7 we see that the baseline B can be divided into two orthogonal dimensions:

$$B_\mathrm{n} = B\sin(90^\circ - \theta + \alpha) = B\cos(\theta - \alpha), \tag{2.35}$$

$$B_\mathrm{p} = B\cos(90^\circ - \theta + \alpha) = B\sin(\theta - \alpha). \tag{2.36}$$

B_p is the projection of B in the radial dimension, it is called *parallel baseline*; B_n is the projection of B in the orthogonal dimension, it is called *perpendicular baseline*. Then the approximation in the equation (2.34) can be viewed as:

$$\Delta\rho \approx B_{\mathrm{p}}. \tag{2.37}$$

This also means the beams from two satellites can be considered as parallel to each other so that the distance difference $\Delta\rho$ is approximated as B_{p} (see figure 2.7). This approximation is also called *far-field approximation*.

3-D Construction

Based on the equations (2.28), (2.34) and (2.36) we have:

$$\phi = \frac{4\pi}{\lambda} B \sin(\theta - \alpha). \tag{2.38}$$

The equation above can also be rewritten as:

$$\theta = \arcsin\left(\frac{\lambda}{4\pi}\frac{\phi}{B}\right) + \alpha. \tag{2.39}$$

The interferometric phase is only defined on the interval $[-\pi, \pi)$, so a *phase unwrapping algorithm* is necessary to solve the ambiguity problem, which is neglected here. Disregarding the ambiguity issue all parameters on the right hand side of the equation above are known, then the look angle is calculable. Using the equation (2.17) the 3-D position of point P is also determinable through polar point determination.

2.5.3 Error budget

Based on the analysis in the last section we can summarize the simplified InSAR processing as following three equations:

$$\phi = \frac{4\pi}{\lambda}\Delta\rho, \tag{2.40}$$

$$\theta = \arcsin\left(\frac{\lambda}{4\pi}\frac{\phi}{B}\right) + \alpha, \tag{2.41}$$

$$H = H_s - \rho\cos\theta. \tag{2.42}$$

The first equation calculates the interferometric phase from the distance difference, the second equation uses this interferometric phase to calculate the look angle, the third equation calculates the height of the target point from the satellite height H_s, the look angle θ and the slant range ρ.

The height of the target point depends on the parameters H_s, ρ, θ, the parameter θ also depends on the parameters α, B, ϕ. The parameter ϕ depends on distance difference $\Delta\rho = \rho_1 - \rho_2$ further, so there is a circular dependence. But the errors in distance measurements are mostly aroused by atmospheric delay and both of the beams are parallel (far-field approximation) and close to each other (small baseline in comparison to large altitude of satellite), so the errors in the

distance difference almost cancel out. So the error in ϕ can be considered as independent of the ρ, and will be treated as an independent stochastic variable.

At the end we have following the functional dependence:

$$H = H\left(H_s, \rho, \theta(\phi, \alpha, B)\right). \tag{2.43}$$

Differentiating the equations (2.40), (2.41), (2.42) we can calculate the height error due to the measurement error of every parameter.

Error from Satellite orbit

$$\delta H = \delta H_s. \tag{2.44}$$

The orbit error is controlled as about 1 cm (Rodriguez, 2008).

Error from slant range

$$\delta H = -\cos(\theta)\delta\rho. \tag{2.45}$$

The error in the slant range comes from electromagnetic (EM) bias effect and path delay in the troposphere and ionosphere. In the SWOT design, the allocated EM bias error is about $1-2$ cm, the allocated error due to ionosphere is about 1 cm, the allocated error due to dry troposphere must be smaller than 1 cm, the allocated error due to wet troposphere is about 1 cm on the ocean, smaller than 5 cm on land (Rodriguez, 2008).

Error from interferometric phase

The equation (2.41) is equivalent to the equation (2.38), differentiating (2.38) with respect to θ and using (2.35) gives:

$$\delta\phi = \frac{4\pi}{\lambda} B\cos(\theta - \alpha)\delta\theta = \frac{4\pi}{\lambda} B_{\mathrm{n}}\delta\theta. \tag{2.46}$$

Differentiating the equation (2.42) with respect to θ gives:

$$\delta H = \rho\sin(\theta)\delta\theta. \tag{2.47}$$

Inserting the equation (2.47) into (2.46) gives:

$$\delta H = \rho\sin\theta\frac{\lambda}{4\pi B_n}\delta\phi. \tag{2.48}$$

Phase error is controlled under 1 cm per 1 km^2 area (Rodriguez, 2008). Dominant phase error sources are:

- Thermal noise in radar signal,
- Decorrelation of the two echoes due to speckle decorrelation of scattered fields,

- Phase imbalance between two interferometric channels.

From equation (2.48) we can also find that error in height estimate is proportional to factor λ/B. If baseline B is reduced, wavelength λ has to be change correspondingly to maintain the same error level, that is also why Ka-band signal is used in SWOT after reduction of baseline to 10 m.

Error from interferometric baseline

When the interferometric baseline B is not fully known, according to the equation (2.38) we have:

$$\delta\phi = \frac{4\pi}{\lambda}\sin(\theta - \alpha)\delta B. \tag{2.49}$$

Inserting this equation into (2.48) gives:

$$\delta H = \rho \sin\theta \sin(\theta - \alpha)\frac{\delta B}{B_n}. \tag{2.50}$$

Error from baseline roll angle

Error in the roll angle α will be interpreted as the error in the interferometric phase ϕ. Differentiating the equation (2.41) with respect to α gives:

$$\delta\theta = \delta\alpha. \tag{2.51}$$

Inserting the equation (2.51) into (2.47) provides:

$$\delta H = \rho \sin(\theta)\delta\alpha. \tag{2.52}$$

According to the equation above an 0.1" error in baseline roll angle results in about 4.5 cm height error at 100 km from the nadir point. The allocated error in the mission design is maximal 2 cm. The error sources are:

- Errors in spacecraft roll estimate,
- Mechanical distortion of the baseline.

A modification of equation (2.52) is that:

$$\delta H = \Delta x \delta\alpha, \tag{2.53}$$

where Δx is the horizontal distance from nadir to target point. Based on this equation the look angle and consequently the swath width is reduced to alleviate the effect of error in roll angle determination.

Chapter 3

Orbit design

3.1 Introduction

For a satellite mission, orbit design is an inevitable question that should be answered. The 3 dimensional motion of satellite in space with respect to the Earth's rotation constrains the function of satellite mission. This relative motion determines geographical coverage, temporal and spatial sampling rates and even the measurement accuracy. These features can be fixed by a few orbital parameters including orbit altitude, inclination, repeat period and eccentricity. Which combination of values is optimal for a satellite mission depends mostly on sensors onboard and scientific applications of the mission.

Since techniques used by SWOT mission are an heritage of the Topex/Poseidon (T/P) and Jason series, and part of their scientific objectives coincide, their orbit selection criteria are similar to some extent. So the first emphasis is put on the review of the orbit design for T/P mission in literature, here the most concerns are the research methodology and constraining conditions, which can help SWOT mission to choose the optimal orbit or some parameters can be directly used with a little modification. The orbit design for T/P mission was comprehensively examined by Parke et al. (1987), their work will be reviewed in the section 3.2. In section 3.3 a realistic application to SWOT is discussed.

3.2 T/P orbit

The orbital dynamics of a satellite and its ground-tracks can be fully determined by the following parameters: satellite altitude, eccentricity of an orbit, inclination of an orbit and repetition period. According to the research by Parke et al. (1987), the altitude for T/P mission should firstly be constrained between 1100 km and 1500 km. The lower limit is determined by orbit perturbations from the atmospheric drag effect, which depends on the gravity of the Earth and sunspot cycle; both of them determine the atmospheric density. The altitude is constrained above 1100 km to decrease the atmospheric friction. The upper limit is determined at the one hand by the complexity of satellite in terms of power and antenna size, and at the other hand by the intensive radiation from Van Allen Belt, which consists of energetic charged particles captured by the Earth's magnetic field from solar wind and cosmic rays. The radiation from Van Allen Belt can degrade the performance of the satellite circuits.

The inclination is constrained by following issues. At first, the T/P mission should cover a possibly large ocean area, including the Drake Passage(62° S). This condition also means the

satellite can cover the Antarctic circumpolar current. According to this scientific purpose of T/P the inclination is supposed to be between 62° and 118° . At the same time the intersection angles between ascending and descending ground-tracks at crossover points are required to be bigger than 40° because a large intersection angle allows the surface currents to be resolved with a similar accuracy in two orthogonal ground-track directions, this condition requires the inclination either smaller than 66° or larger than 107° as calculated by Parke et al. (1987). The repetition period should be between 10 and 20 days in order to investigate the ocean dynamic topography.

Meanwhile the tidal effect also cause the ocean surface derive from the ocean dynamic topography, this effect should be removed from the measurement values. The most largest tidal constituents are frequency signals compared with all possible sampling periods for T/P mission, so the Nyquist sampling frequency can not be fulfilled, consequently in the spectrum domain the high frequency signals (tides) will be aliased into the base band (geoid and annual cycles) and mixed with low frequency signals there, so generally both of them can not separated from each other. However Parke et al. (1987) pointed out, these tidal signals can be resolved from the base band signals if "the aliased frequencies are sufficiently different from zero or one cycle per year", so that "data can be combined spatially to provide enough phase information to solve for the tide over a period of years, provided that the orbit is not nearly sun-synchronous."

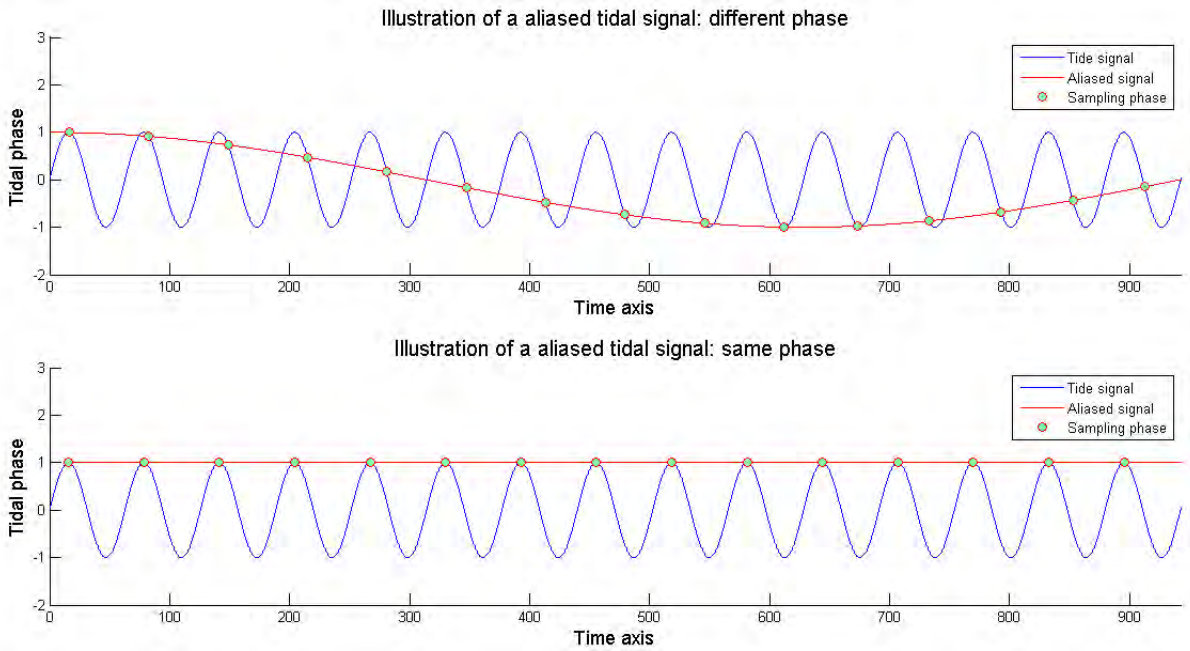


Figure 3.1: *At the top: Repetition period is not a multiple of tidal period, different phase will be sampled in advance. At the bottom: Repeat period is a multiple of tidal period, same phase will be sampled.*

An intuitive explanation about this viewpoint can be found in a remote sensing book by Rees (2001). Assuming that the repetition period of the satellite is 3.01 days, and period of a tidal constituent is 0.5 day, so after one sample interval the tide has occurred 6.02 times and measured by the satellite with 0.02 period in advance, so that after 150.5 days all tidal phases will be sampled and can be resolved from measurement values. But if the the satellite period is an exact multiple of the tidal period, e.g. 3 days in this example, the phase measured by the

satellite will be all the same and no tidal constituents separable. An illustration is given in the figure above.

Parke et al. (1987) also provided an explanation for the issue of sun-synchronous orbits. An important fact for synchronous orbit is that the orbit plane rotates and keeps a constant angle with respect to the solar direction, consequently the satellite always crosses a given latitude at the same local time. This implies that the satellite ground tracks repeat at the same local time, hence the S2 tide will be measured only in a same phase. This situation is illustrated in the figure below, the geometrical configuration among sun, satellite and local point is always same at the time of satellite's overpass.

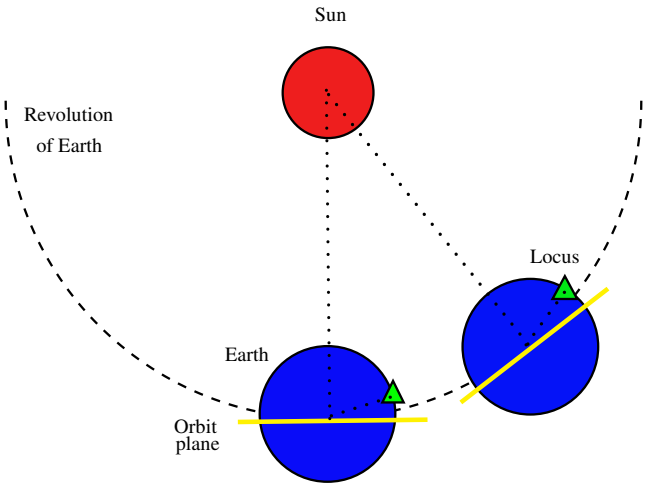


Figure 3.2: *Problem of sun-synchronous orbit*

Parke et al. (1987) proposed two requirements for the eight largest tidal alias frequencies, so that on the one hand the aliased frequencies would not be mixed by the ground frequency, on the other hand the aliased frequencies would not overlap and could be separated from each other. The first one was nonzero condition: the aliased tidal frequencies must be at least two cycles per year. The second one was separability condition: the aliased frequencies must differ by at least one complete cycle over the length of the T/P mission. Based on the consideration above, the inclination is constrained under 66° and the repetition period must less than 7 days or exactly equal to 10 days.

The last important issue is the condition for repeat orbits, which is preferable because of its successive data acquirement over the same ground-track and equally spaced sampling in time and ground-tracks. The intersection of all above conditions called "Topex/Poseidon window" by Parke et al. (1987) leading to:

- altitude: 1100 km - 1500 km,
- inclination: $62° - 66°$,
- repetition period: 10 days .

Finally some secondary requirements are taken into account and orbit parameters are fixed as 1335 km for altitude, 64.8° for inclination and 10 nodal days for repetition period. This type of configuration has proved to be a great success during the T/P mission and it has also been adopted by the successive altimeter missions such as Jason series.

3.3 SWOT orbit

3.3.1 Review of T/P orbit

Orbit study by Parke et al. (1987) for T/P provides us a collection of constraints that should be concerned during orbit design. They are summarized in table 3.1.

Effect	Constraint
Atmospheric drag	$H > 1100$ km
Radiation from Van Allen Belt	$H < 1500$ km
Intersection angles at crossover	$i < 66^\circ$
Geographical coverage	$i > 62^\circ$
Tidal aliasing effect	$T = 10$ days

Table 3.1: *Constraints for T/P*

These considerations must be modified to reflect the actual requirements for SWOT mission. In the next some analysis in the literature will be introduced.

3.3.2 Starting point of analysis

Stammer and Siegismund (n.d.) provided an analysis about the orbit optimization for EUMETSAT (European Organization for the Exploitation of Meteorological Satellites) altimeter missions. Here some information about SWOT orbit can be found. They also provided a general discussion about orbit design for the future altimeter missions. In comparison with the last successful T/P orbit, the alternative orbits in the future should have following characters:

- Altitudes below 1000 km,
- Inclinations near 78°,
- Repeat periods between 9 and 21 days.

These recommendations are also useful for SWOT mission, and they can be considered as a starting point of the most orbit design for the future altimeter missions. The first term is to reduce radiation exposure of the satellite in space, it will also decrease the technical complexity and consequently the mission costs. The second one is to improve latitudinal coverage. The third one is to improve spatial resolution.

In the meeting "Mesoscale and Submesoscale Oceanic Processes: Explorations with Wide-Swath Interferometry Radar Altimetry" held in April 2008 at Scripps Institution of Oceanography, Steven Nerem et al. (2008) made a presentation about orbit design for SWOT from oceanographic community. The starting point of their work is following:

- Altitude: $800 - 1000$ km,
- Repeat period: $15 - 25$ days,
- Inclination: 78°,

A comprehensive simulation and analysis can be found in the script of his presentation, which will be presented in the next section.

3.3.3 SWOT orbit design by (Nerem et al., 2008)

According to (Steven Nerem et al., 2008), the SWOT orbit is determined from following aspects:

- Latitude coverage,
- Spatial sampling,
- Tidal aliasing,
- Temporal sampling.

We will discuss them in details respectively.

Latitude coverage

As SWOT will observe both of ocean and terrestrial water, so the orbital inclination has to be increased to improve latitude coverage so that all important rivers can be observed. For SWOT the candidate inclination is 78°.

Spatial sampling

As we know a great drawback for T/P orbit is the large equatorial spacing about 300 km. This coarse spatial resolution prevents a further study of submesoscale processes in the ocean. In order to improve spatial sampling rates the repeat period should be increased to allow more orbit revolutions in a period, and consequently a more denser sampling on the equator.

The footprint of SWOT ground-track is about 130 km, so the equator (perimeter of Earth is about 40000 km) can be fully sampled by about 300 orbital revolutions, this number also constrains the repeat period. As simulated by Steven Nerem et al. (2008), the repeat period should be greater than 22 days to fully sample the equator.

An other important concern is about subcycle determination. A subcycle is similar like a near repeat orbit in a relative larger tolerance. Fu and Shum (2008) explained why this issue was important for SWOT. The SWOT is expected to study high frequency temporal variability of submesoscale processes for a limited duration, during this time the repeat period should be reduced to achieve higher sampling rates, this study is important for coastal zones. The idea of repeat period adjustment is accomplished by changing orbital altitude. Fu and Shum (2008) pointed out that a 22-day orbit with a 3-day subcyle could be easily adjusted from an altitude of 970 km to 976 km to get into a 3-day repeat orbit. This one is also the final design for SWOT as seen in the next. In general an adequate subcycle is desired to perform repeat period adjustment as easily as possible, which also constrains the choice of repeat period. Steven Nerem et al. (2008) simulated the possible repeat periods and altitudes under 78° inclination, they only considered the candidates with short subcycle ($3-5$ days). With the condition of repeat orbit

there are not many combinations between repeat period and altitude, only 19 combinations are taken into account for further analysis.

Steven Nerem et al. (2008) also provided some visual simulations of subcyle for different repeat period. A method used by Parke et al. (1987) can be applied here to assess the subcycle period given by Steven Nerem et al. (2008). An example is give in the following for a 22-day orbit with 301 orbit revolutions in a nodal day. In this term thesis a Matlab program is designed to simulate the orbit ground-track and visualize the explanation in the following. The Matlab code can be found in the appendix.

At first we could calculate the number of orbit revolutions per nodal day:

$$\frac{301}{22} = 13 + \frac{15}{22}. \tag{3.1}$$

This analytical result can also be verified by the following simulation result:

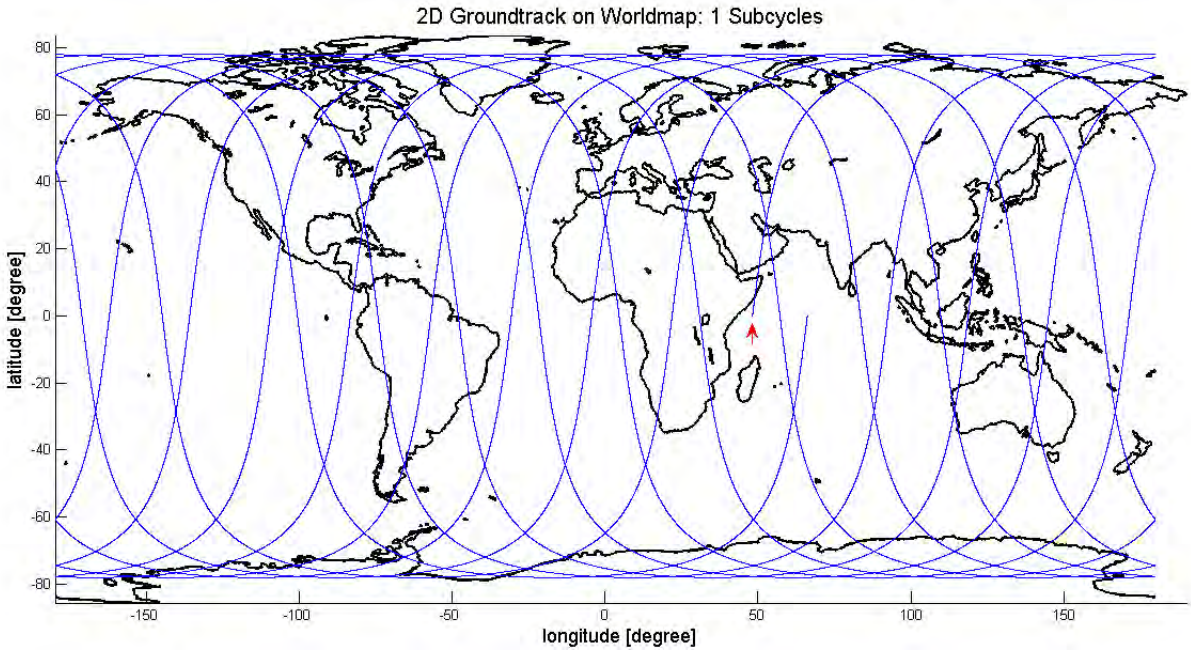


Figure 3.3: *Revolutions per nodal day*

The figure above is the ground-track of the SWOT satellite after nearly one nodal day (13 revolutions) from simulation. We can see that after 13 revolutions the orbit does not coincide with its starting point, which is marked by the red arrow near equator. The fractional part 15/22 is the reason why the orbit does not repeat in one nodal day, so 22 days are needed to conquer this fractional part.

We can also see how this fractional part will be conquered through other simulation results:

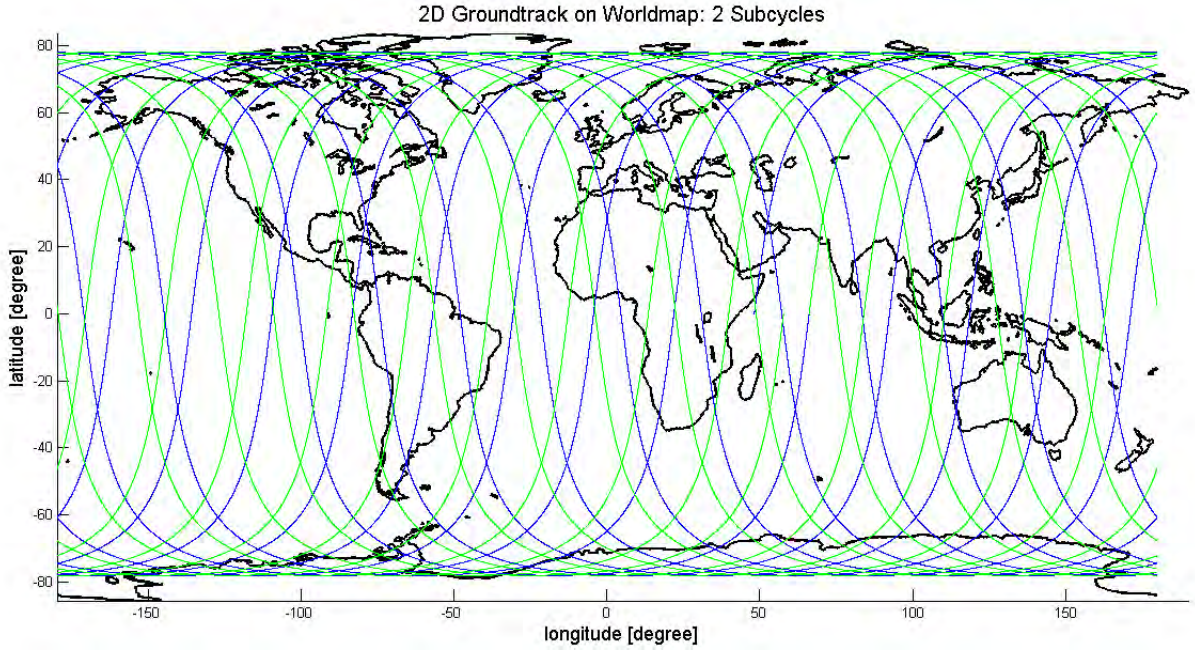


Figure 3.4: *Ground-track after 2 nodal days*

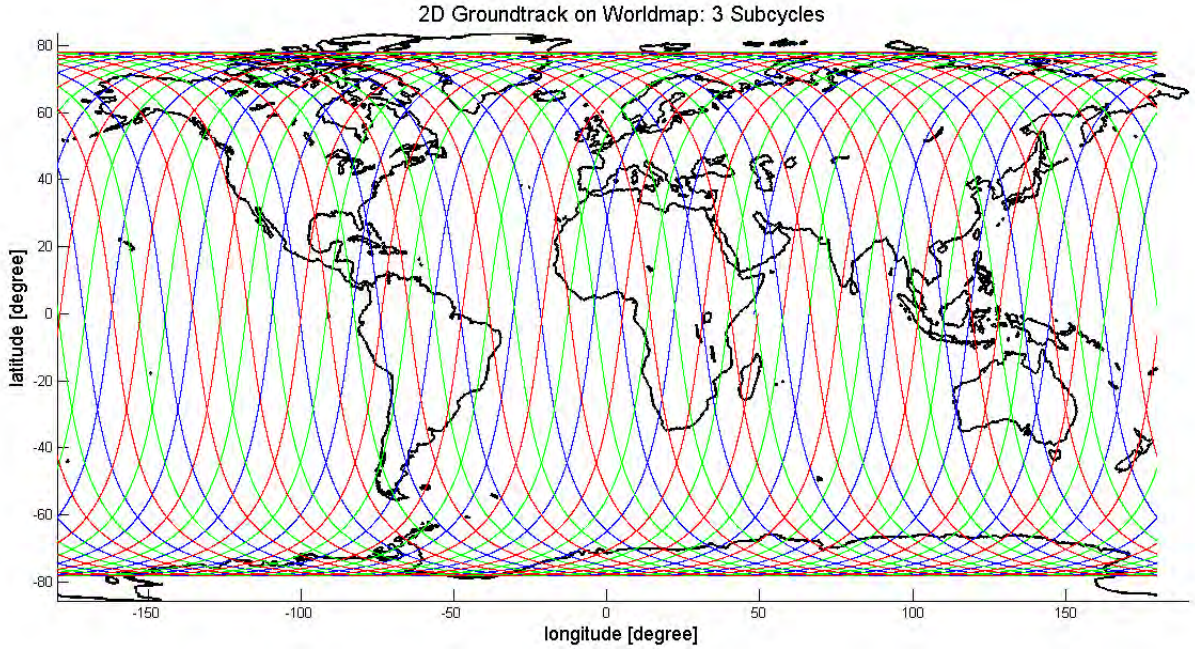


Figure 3.5: *Ground-track after 3 nodal days*

In the figure (3.4) the ground-track for the first nodal day is plotted in blue color, whereas the ground-track for the second nodal day is plotted in green color. We can see that the green line can be considered as a duplicate of the blue one only shifted westward. In the figure (3.5) the ground-track for the third nodal day is plotted in red color. It is also a westward duplicate of the second ground-track in green color.

At the same time the Earth's surface looks like being homogeneously and completely covered by these three ground-tracks. The orbit seems to be repeating after three nodal days. This situation can be also analytical proved:

After 3 nodal days, there will be so many revolutions repeated by the satellite:

$$3 \times \frac{301}{22} = 3 \times (13 + \frac{15}{22}) = 41 + \frac{1}{22}. \tag{3.2}$$

1/22 is the the smallest possible derivation from the exact repeat period (0/22). If this relative larger tolerance is allowed the orbit is nearly considered to be repeating in a cycle of 3 days. This fractional part 1/22 also means so many ground-track spacings, so we also could calculate this tolerance in the following way:

$$\frac{40000}{13 + \frac{15}{22}} \times \frac{1}{22} = \frac{40000}{301} \approx 133, \tag{3.3}$$

so the tolerance is about 133 km.

In order to visualize this analytical result, the ground-track for the forth nodal day is added:

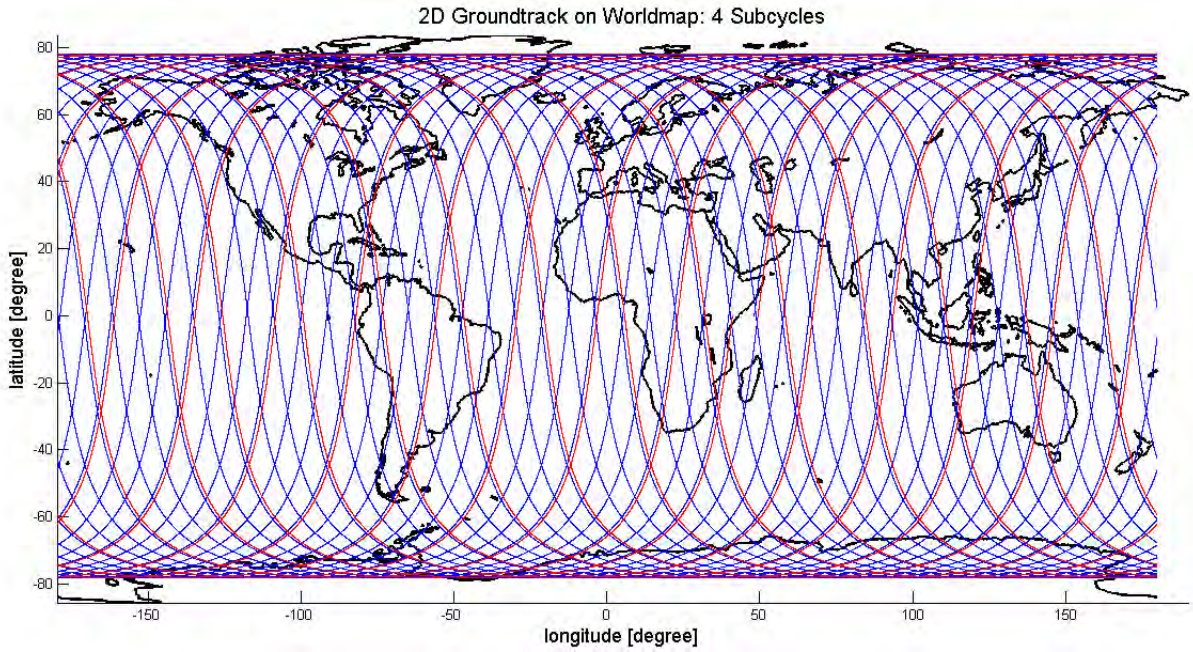


Figure 3.6: *Ground-track after 4 nodal days*

The preceding three ground-tracks are plotted in blue color, the forth one is plotted in red. We can see that the red line nearly coincides with the blue one only with a tiny displacement. Figure (3.6) can also be amplified to estimate this tiny displace visually:

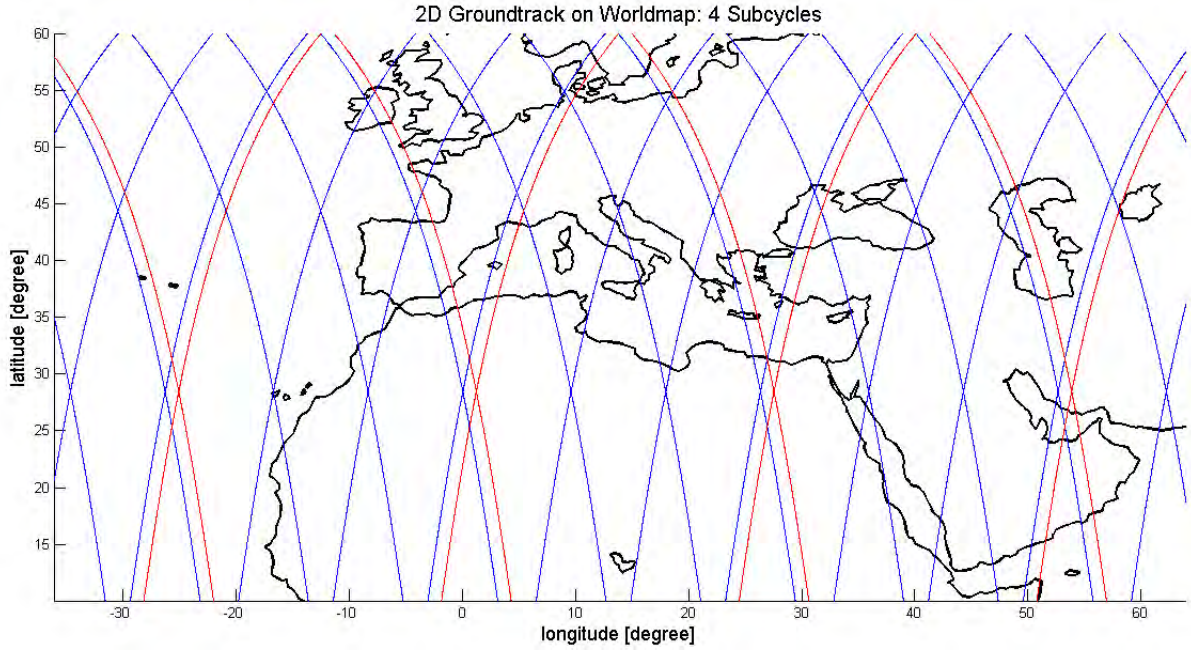


Figure 3.7: *Clipping of ground-track after 4 nodal days*

We can see that this tiny displace near the equator is about 1° (1° corresponds to 111 km approximately), this is consistent with the analytical result from equation (3.3).

Two further figures are presented here to visualize the repeating orbit of SWOT satellite:

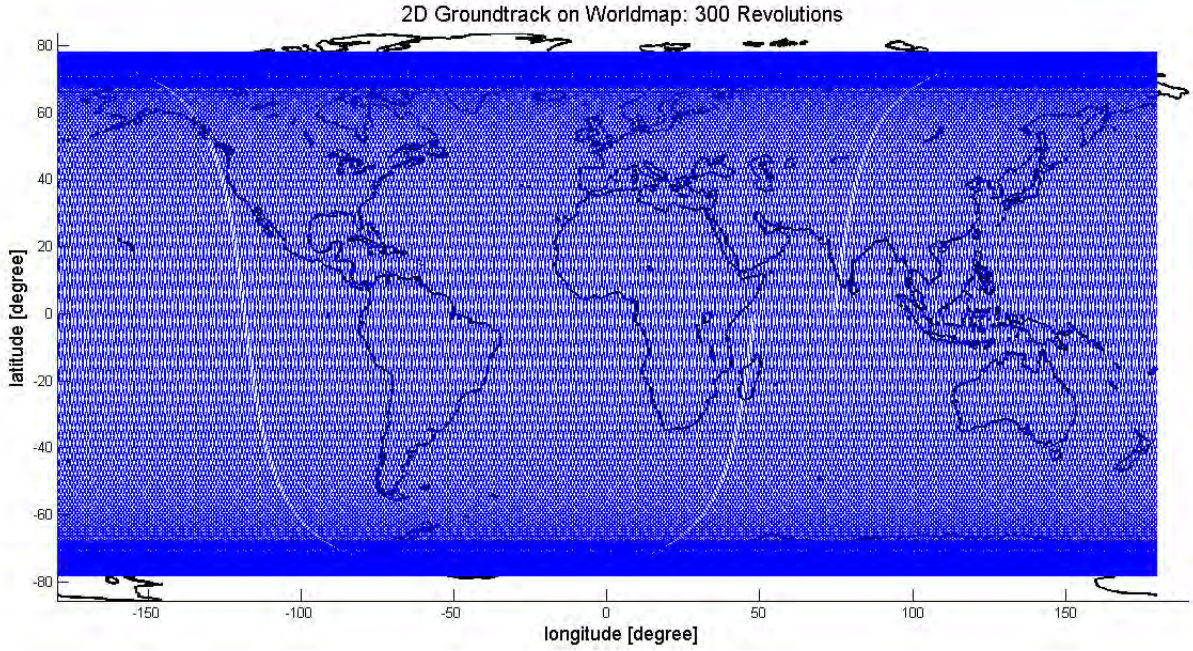


Figure 3.8: *Ground-track after 300 revolutions*

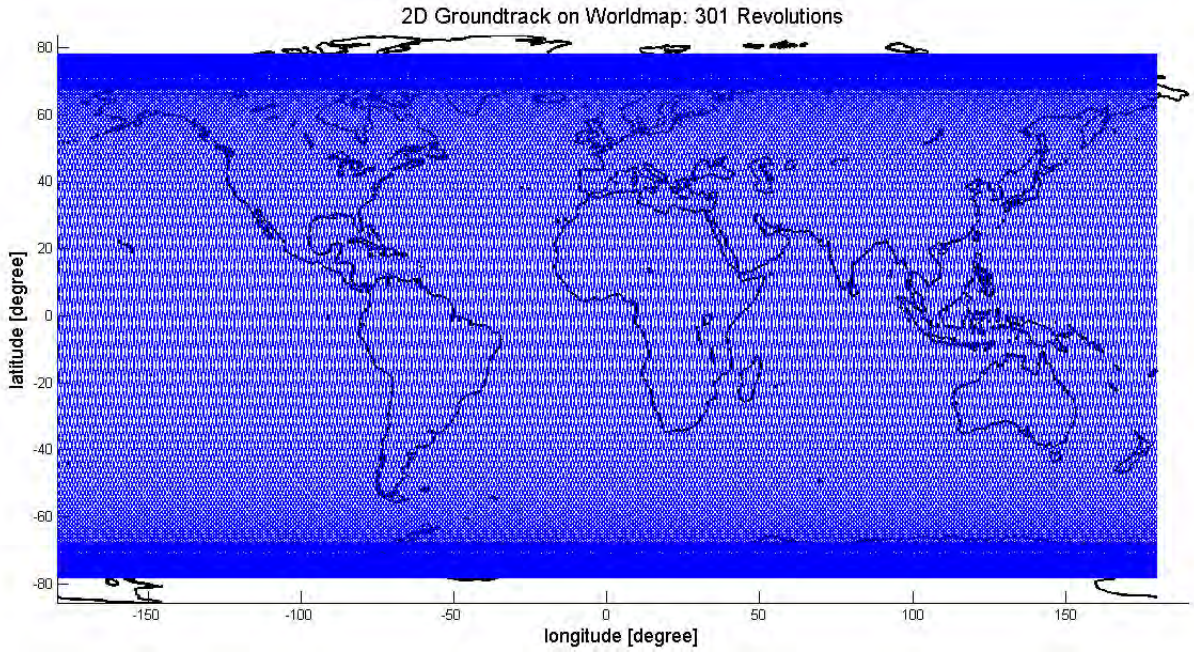


Figure 3.9: *Ground-track after 301 revolutions*

In the figure (3.8) we see that the whole Earth's surface is nearly homogeneously sampled only with a gap, which is in the figure (3.9) fulfilled. The whole ground-track is then closed by the last revolution.

The following figure provides an amplified version of the figure (3.9). We can see that the the whole ground-track is closed.

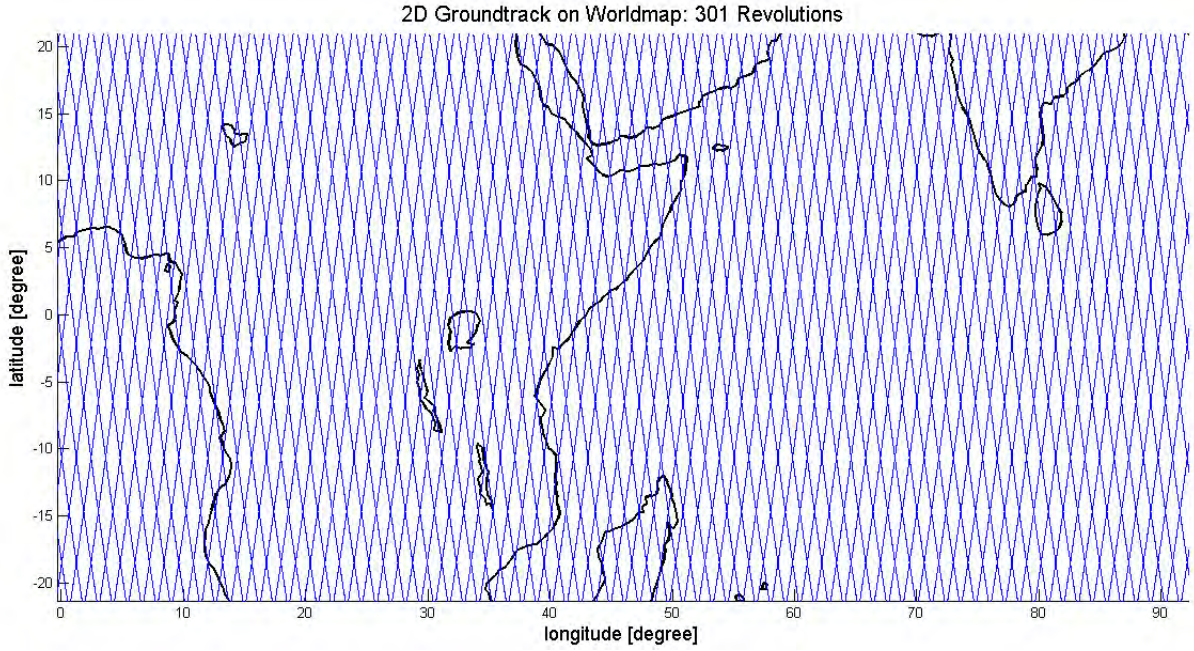


Figure 3.10: *Clipping of ground-track after 301 revolutions*

Tidal aliasing

Similar to the discussion by Parke et al. (1987) for T/P orbit, the tidal aliasing effect is also a great concern. Steven Nerem et al. (2008) simulated the aliased frequencies with different altitude and inclination values, the selection criteria is also similar to that used by Parke et al. (1987):

- Good separation of major tide constituents aliasing frequencies,
- Alias frequencies should not be close to one cycle per year,
- Tides should not alias to very long periods($\ll$ 1 year).

According to these criteria there are only 3 candidates of repeat period under the inclination of 78°: 15, 21, 22 days. The 22-day repeat period is preferred because of its shorter subcycle (3 days), for 15-,21-day repeat period the subcycle are both 4 days. The corresponding altitude for 22-day period is 970.6 km. Steven Nerem et al. (2008) also simulated sampling of tides by SWOT under the 22 day repeat period and discussed the effect of additional sampling on solution of tidal aliasing. A drawback of this choice was that the aliased frequency of K1 tidal constituent would be obviously smaller than 2 cycle per year (1.3 cycle per year), this offended the second criteria.

Temporal sampling

Steven Nerem et al. (2008) also provided a temporal sampling analysis for different latitudes: Agulhas current region at mid-latitude, Amazon river near equator and Lena river at high-latitude. Plots of revisit number is also given by them for 4-,10- and 22-day sampling periods. The result is region and latitude dependent. In general, the number of observations at high latitude will be larger than that near Equator, and at crossover points there will be more observations than that in the surrounding areas. In the table 3.2 a summary is given after the work by Steven Nerem et al. (2008).

Sampling	**Agulhas Current Region**		**Amazon River**		**Lena River**	
period [days]	**Latitude [°]**	**Nr/cycle**[1]	**Latitude [°]**	**Nr/cycle**	**Latitude [°]**	**Nr/cycle**
22	30	2 – 3	0	2	60	4 – 5
	60	5 – 6	15	2 – 3	75	10
10	30	1 – 2	0	2	60	3 – 4
	60	4 – 5	15	1	75	7 – 8
4	30	1	0	2	60	2 – 3
	60	2	15	1	75	3 – 4

[1] **Nr/cycle**: Sampling number per cycle.

Table 3.2: *Temporal sampling at different latitudes*

3.3.4 Consideration from hydrology

As SWOT will play an important role in monitoring the storage of terrestrial water bodies, some scientists also consider the requirement of the orbit configuration for hydrological applications.

The 78° inclination is mostly due to the desire from hydrology. For hydrology tidal aliasing issues are not important because terrestrial water do not have elevations that systematically vary on such short time scales. More important is the revisit number per cycle.

Lee et al. (2010) provided a sensitivity study of storage change accuracy to SWOT orbital parameters. They chose 15-, 21-, 22-day (these are candidates in the work by Steven Nerem et al. (2008), see section 3.3.3) as repeat period candidates, 74°, 78° as inclination candidates, and simulated the storage change for lakes in the Peace-Athabasca Delta in Canada, Northern Alaskan and West Siberian Arctic Regions using synthetic SWOT measurements. According to their work, there is no preferential choice among the potential SWOT orbit, storage change errors are not sensitive to the orbital configuration but the shape of the water body is more important in estimating storage change using SWOT measurements.

3.3.5 Final orbit parameters

Here is the final orbit parameters for SWOT, cited from Stammer and Siegismund (n.d.)

	Orbit parameters	**SWOT**
Independent	Orbits per repeat	301
	Exact repeat period [days]	22
	Inclination [°]	78
Derived	Revolutions per day	$13 + 15/22$
	Altitude [km]	970.6
	Equatorial spacing [km]	133
	Subcycle period [days]	3

Table 3.3: *Final orbit design for SWOT*

Chapter 4

Radiometer and its application

4.1 Introduction

The atmosphere around the Earth consists of a mixture of gases and other liquid and solid particles. Due to these constituents most harmful radiations from cosmic will be filtered out and the temperature variation on the Earth is also reduced, so that lives on the Earth are protected. On the other hand this filtering function of atmosphere also prevents the transparency of radiation at special wavelengths. These spectral regions at which the atmosphere is transparent, are also known as the *Atmospheric Windows*. For remote sensing, the interaction of radiations at these regions with atmosphere are investigated and used to sound the atmospheric information.

Historically the visible and infrared bands are firstly used to observe the Earth's surface and atmosphere, then followed by the use of microwave regions. The major problem for soundings at infrared and visible bands is that they can not used in cloud conditions, in order to acquire information of atmosphere under clouds the microwave band will be used.

The transmittance of atmosphere in the microwave regions is not uniform mostly due to the varying amount of its absorbing constituents: water vapor and oxygen, water vapor has resonant absorption at 22 and 183 GHz, and the strongest absorption due to oxygen occurs at 118 and between 50 and 70 GHz. Another varying constituent in the atmosphere is liquid water clouds. The microwave absorption by liquid water clouds increases monotonically with frequency and is minimal at frequencies below 20 GHz.

These characteristics of clouds at microwave frequencies are often used to determine the atmosphere status. Radiometric measurements in the 50-60 GHz are used to derive temperature profiles in clear and cloudy atmospheres, measurements at 22 GHz are used to determine column water abundance, measurements around 183 GHz are used to obtain humidity profiles. At the frequencies between 20 and 40 GHz, absorption by clouds and larger liquid drops in rain can be used to derive information about water content and horizontal extent.

In the history, many satellite microwave radiometers have been launched to the space to observe the Earth's surface and atmosphere. they have been used to measure rainfall, cloud liquid water, water vapor and the vertical distribution of temperature. An overview of these missions can be found in the literature (Grody, 1993) and (Gentemann et al., 2010).

Besides these satellite microwave radiometers designed for meteorological, hydrological and oceanographic applications there are also other types of radiometers, for examples, special-purpose radiometers for the correction of radar altimeters for tropospheric path delay.

Altimeter is an instrument used to measure the altitude of an object above a fixed level. The principle used by altimeter is a timing system based on the assumption of constant wave velocity the distance between satellite and ocean surface can be accurately calculated. However as electromagnetic signals pass through the a medium. The refraction index will be changed and and consequently the wave velocity. So there will be a path delay as the signal go through the atmosphere. The error comes from two sources: ionosphere and troposphere.

The ionosphere is full of ions because of intensive radiation from the Sun. These electrons make the ionosphere act as a dispersive medium, in which the group velocity of a electromagnetic signal can be changed, and consequently the distance measurement could be incorrect. The error made by the ionosphere can be calculated from a parameter called *total ionospheric content*. The troposphere involves many gas molecules and liquid, solid particles which can delay the propagation of signal in it. For altimeter working on microwave bands, the largest influence comes from water vapor content along the signal path. Whether total ionospheric content or water vapor content has more effect on the signal delay, this depends on the signal frequency. In general, the larger the frequency, the smaller the delay effect from ionosphere but larger from troposphere,and vise versa. So for the L-band altimeter, the ionospheric delay is more critical, but for Ka-band altimeter, just the case for the SWOT mission, the most concern is water vapor content.

The passive microwave radiometer for water vapor correction is not an innovative in SWOT, but inversely SWOT inherits this principle from previous altimeter mission, for example T/P, Jason-1/-2. All of these radiometers are designed by JPL. Even before T/P, there were many satellite mission using microwave radiometer to correct their distance measurements from space, for example Special Sensor Microwave/Imager (SSM/I) on board the United States Air Force Defense Meteorological Satellite Program (DMSP) was used to retrieve meteorological parameters over the ocean. The working frequencies are similar to that used in T/P and SWOT. The table below compares the the working frequencies of the radiometer used by SSM/I,T/P and Jason-2.

Mission	Radiometer	1.frequency	2.frequency	3.frequency	4.frequency
DMSP	SSM/I	19.4 GHz	22.2 GHz	37 GHz	85.5 GHz
T/P	TMR	18 GHz	21 GHz	37 GHz	-
Jason-2	AMR	18.7 GHz	23.8 GHz	34 GHz	-

Table 4.1: *Operating channels of different radiometers*

We can see from this table that the operating channels of these radiometers has not yet changed so much. SWOT will also use three frequencies between $18 - 37$ GHz to correct the range delay over ocean. In addition to them high frequencies above 90 GHz will also be used to perform range correction on land. Discarding this change, the physical background and retrieval algorithm of these radiometers are almost similar. So during the literature research some descriptions about T/P radiometer and SSM/I will also be presented, this context provide us a better access to the algorithm used for path delay correction in SWOT.

4.2 Physical background

4.2.1 Radiative transfer theory

The physical principle for the path delay correction is based on the *radiative transfer theory* for a non-scattering thermal medium. These theory models how the energy variation behaves as a radiation specially migrate from one place to another. As electromagnetic waves transmit through the atmosphere, it also leads the energy variation of surrounding medium. This phenomenon can be great modeled by radiative transfer theory. An analytical derivations are given in the appendix, cited from (Grody, 1993). The main result is as follows:

$$T_{\mathrm{b}}(\nu) = T_{\mathrm{b0}}(\nu)e^{-\tau(s_0)} + \int_0^{s_0} T(s)\alpha e^{-\tau(s)}ds. \tag{4.1}$$

The parameters in this equation is as follows:

- wave frequency: ν,
- absorption coefficient of medium: α,
- opacity or optical thickness τ, which determines the extent of attenuation effect by the medium. This term is defined as :

$$\tau(z) = \int_z^H \alpha(z')dz', \tag{4.2}$$

- distance between observer and background: s_0,
- measured brightness temperature by observer: T_{b},
- brightness temperature from background radiation: T_{b0}.

This equation has a clear physical meaning. It means that the measured brightness temperature can be divided into two parts, as described at the right-hand side of this equation. The first one comes from the background radiation, characterized as T_{b0} and the exponential term $e^{-\tau(s_0)}$ describes attenuation effect during the propagation through medium. It is a function of distance between the background and observer. The second term is a summation of radiation energies from medium as they propagate from different height s and accordingly attenuated by a factor $e^{-\tau(s)}$.

4.2.2 Application in meteorology

(This section is referenced from the work by (Janssen, 1993))

The equation (4.1) can be applied to the satellite case. Under the assumption that the radiation propagates from the Earth's surface to the satellite height in the normal direction, and here the

background is the Earth's surface and observer is the satellite radiometer, so the equation (4.1) can be rewritten as:

$$T_b(\nu) = T_u(\nu) + T_e(\nu)e^{-\tau(0)} + T_r(\nu)e^{-\tau(0)}. \tag{4.3}$$

The left hand side of this equation is the measured brightness temperature by radiometer in space. Three contributions determine the measurement value, as listed on the right hand side of the equation above:

1. upwelling radiations of atmosphere T_u,
2. radiations emitted by ocean surface T_e,
3. reflected atmospheric radiation on the ocean surface T_r.

T_u, T_e and T_r are explicitly formulated as following:

$$T_u(\nu) = \int_0^H T(z)\alpha e^{-\tau(z)}dz, \tag{4.4}$$

$$T_e(\nu) = \varepsilon(\nu)T_s, \tag{4.5}$$

$$T_r(\nu) = (1-\varepsilon(\nu))\left[\int_0^H T(z)\alpha(z)e^{-\int_0^z \alpha(z')dz'}dz + T_c(\nu)e^{-\tau(\nu)}\right]. \tag{4.6}$$

where $T(z)$ is temperature profile of atmosphere, T_s is the surface temperature, T_c is the cosmic background flux, which is a constant. α is atmospheric absorption coefficient, $\varepsilon(\nu)$ is the effective emissivity of the ocean surface.

The bracket in equation (4.6) describes the sum of the radiation from atmosphere downloading effect and cosmic background. They are both reflected by the ocean surface and then attenuated by the atmosphere before they received by the sensors. The term $(1-\varepsilon(\nu))$ comes from the Kirchoff's law. Because the entire energy received by the surface can either be absorbed or reflected, if the reflectivity is $(1-\varepsilon(\nu))$, so the absorbency must be $\varepsilon(\nu)$. According to the energy conservation the absorbed energy of the material is finally emitted so the emittance is also $\varepsilon(\nu)$.

The equation (4.3) is the fundamental for atmospheric soundings at microwave bands. Many objects of scientific study are included in this equation, such as surface temperature T_s, emissivity of ocean surface $\varepsilon(\nu)$, temperature profile $T(z)$ or atmospheric status characterized by opacity τ, which depends on the wave frequency. An application of this frequency dependence is that different wavebands within the atmospheric windows are combined to sound atmosphere, especially at the wavelengths where the atmospheric attenuation effect is significant. For the purpose of retrieval of atmospheric water vapor contents the operating frequencies near 22 GHz is usually used just as mentioned before. This spectral region is also called *(weak) water absorption line*. A few frequencies near the swing of water absorption line will also be elected to compensated other perturbing effect due to liquid water and variation of ocean surface. (see table (4.1))

It is also worth noting that equations (4.1) and (4.3) are equivalent to each other. Equation (4.3) is a general formula in physics and mostly used for theoretical analysis, whereas equation (4.3) provides more clear interpretations in meteorology. Equation (4.3) also give us a qualitative insight into the relationship between measured brightness temperature and relevant parameters. If equations from (4.3) to (4.6) are combined, we see that the brightness temperature is

actually a function of $T(z)$, $\alpha(z)$, T_s and ε. Furthermore the atmospheric absorption coefficient can be determined through atmospheric pressure profile $P(z)$, temperature profile $T(z)$, density of water vapor $\rho_v(z)$ and density of liquid water $\rho_l(z)$, so finally the measured brightness temperature is a function of 6 parameters (neglect frequency ν):

$$T_b = T_b(T_s, \varepsilon, T(z), P(z), \rho_v(z), \rho_l(z)). \tag{4.7}$$

The first two parameters relate to the surface properties, the others are atmospheric parameters. Equation (4.3) is usually used in a numerical model to simulate measurement values. Correspondingly equation (4.7) provides a qualitative analysis (see section 4.3 and 4.4).

4.2.3 Relation between water vapor and brightness temperature

An application of equation (4.1) is to find relationship between measured brightness temperature and total water vapor contents. In literature Wei et al. (1989) provided a good review of retrieval algorithms for water vapor contents, and an explanation for a multi-frequency method can be found there. The equation 4.1 can be integrated by introducing a mean radiating temperature of the atmosphere, T_m, so that

$$T_b(\nu) = T_{b0}(\nu)e^{-\tau(s_0)} + T_m \int_0^{s_0} \alpha e^{-\tau(s)} ds. \tag{4.8}$$

Because α is approximately 1, the solution of this equation is given by them as:

$$T_b(\nu) = T_{b0}(\nu)e^{-\tau(s_0)} + T_m(1 - e^{-\tau(s_0)}), \tag{4.9}$$

where s_0 is approximately supposed to be infinity. Thus from equation above the opacity can then be calculated as:

$$\tau(s_0) = -ln(\frac{T_m - T_b}{T_m - T_{b0}}). \tag{4.10}$$

The opacity of atmosphere is mostly due to water vapor, oxygen and liquid water, so the Wei et al. (1989) rewrote the opacity as follows:

$$\tau(s_0) = \tau_v + \tau_o + \tau_l, \tag{4.11}$$

where τ_v, τ_o, τ_l denote opacity due to water vapor, oxygen and liquid water respectively. These terms can be explicitly formulated as (here $s_0 = \infty$):

$$\tau_v = \int_0^\infty \rho_v k_v dz = K_v V, \tag{4.12}$$

$$\tau_l = \int_0^\infty M k_l dz = K_l L, \tag{4.13}$$

$$\tau_o = \int_0^\infty \rho_o k_o dz, \tag{4.14}$$

where ρ_v, ρ_o are the densities of water vapor and oxygen, M is the liquid water content (the mass of water per unit volume of air), k_v, k_o, k_l are the mass absorption coefficients respectively, K_v, K_l are weighted averages of k_v, k_l (oxygen is homogeneous mixed in the atmosphere so

there is no weighted average term for k_o). At last, V, L are the total water vapor and total liquid water in the atmosphere.

Based on the equations (4.10), (4.10), (4.12) it is possible to relate the brightness temperature to total water vapor and total liquid water. If a two-frequency approach is used, then from equation (4.10), (4.12) we have:

$$\tau_1 = K_{v1}V + K_{l1}L + \tau_{o1}, \tag{4.15}$$

$$\tau_2 = K_{v2}V + K_{l2}L + \tau_{o2}. \tag{4.16}$$

So L, V can be solve from τ_1, τ_2:

$$L = \frac{f_2 K_{v1} - f_1 K_{v2}}{K_{v1}K_{l2} - K_{v2}K_{l1}}, \tag{4.17}$$

$$V = \frac{f_2 K_{v1} - f_1 K_{v2}}{K_{v1}K_{l2} - K_{v2}K_{l1}}. \tag{4.18}$$

with equation $f_1 = \tau_1 - \tau_{o1}$ and $f_2 = \tau_2 - \tau_{o2}$.

Because opacity is related to the measured brightness temperature by equation (4.10), so finally the total water vapor content is related to the brightness temperature measured at two different operating frequencies. Although there are some approximations during derivation, this result still indicates an approach to retrieve the total water vapor through a combination of brightness temperature measurements at different frequencies. In general more than two frequencies can be used.

4.3 A statistical approach

As shown in the last section, the measured brightness temperature is related to the total water vapor content. This relationship should be explicitly formulated as a function, but the analysis given in the last section uses many assumptions and is not accurate enough. A commonly used approach to determine this function is a statistical approach. The idea in background is to use physical models and radiosonde records to calculate the expected brightness temperatures at the frequencies of measurements, then this simulated data set will be used to derive statistical relationships (least square technique, regression technique) between water vapor content and measured brightness temperatures.

Schluessel and William (1990) provided a detailed description of multichannel (19, 22, 37, 86 GHz) statistical approach for water vapor content retrieval by SSM/I measurements. In the following this example will be briefly presented in order to obtain a better insight into this technique. Their work was also cited and modified by Keihm et al. (1995) during design of a pre-launch atmospheric correction algorithm for T/P mission.

Schluessel and William (1990) have firstly simulated measurement value of the brightness temperature from SSM/I sensor using the radiative transfer equation, in which the brightness temperature was derivable from following parameters: atmospheric transmission coefficient and emissivity, water vapor and liquid water profile in the atmosphere, temperature of ocean surface, salinity of ocean and wind-stress on the ocean (see 4.7). Schluessel and William (1990)

used (Liebe, 1985) model to calculate the atmospheric transmission coefficient. The atmospheric emissivity was divided into two parts: one was the emissivity from plane surface, the other was from rough surface, which depended on wind speed on the ocean. These two parts had already been modeled by Klein and Swift (1977) and Wentz (1983). The other parameter used to derive brightness temperature were then randomly generated based on empirical measurement values. These enclosed a large set of oceanic and atmospheric situations. At last the instrument noise was also added to the simulation.

Based on the measurement value of the brightness temperature simulated by the approach above, Schluessel and William (1990) found that the brightness temperature simulated in all channels (here 19, 22, 37, 86 GHz) was highly correlated with integrated water vapor contents, especially in 19 and 22 GHz channels. The correlation between simulated brightness temperature and water vapor contents were almost linear according to the figures in their paper. According to these results a multidimensional linear function was plausible and natural to model these relationship. However, Schluessel and William (1990) hat modified this linear function as a result of following considerations: The water vapor content was dominate by 22 GHz channel because it was near 22.2 GHz water vapor line. The linear relationship between 22 GHz brightness temperature and water vapor contents was only valid until about $3\mathrm{g/cm^2}$ of water vapor contents. And the study by Chang and Wilheit (1979) indicated that a logarithmic function was more adequate to fit the relation for 22 GHz, so the linear relationship for 22 GHz was modified to $ln(280 - T_{\mathrm{B}}(22\mathrm{GHz}))$. Schluessel and William (1990) finally established a functional model between water vapor contents and brightness temperature in different combinations of frequencies.

$$V = a_0 + a_1 ln(280 - T_{22}), \tag{4.19}$$

$$V = b_0 + b_1 ln(280 - T_{22}) + b_2 \left[T_{19} - ln(280 - T_{22})\right], \tag{4.20}$$

$$V = c_0 + c_1 ln(280 - T_{22}) + c_2 \left[ln(280 - T_{22}) - T_{37}\right], \tag{4.21}$$

$$V = d_0 + d_1 ln(280 - T_{22}) + d_2 (T_{37} - T_{86}^{0.8}), \tag{4.22}$$

where V denoted the water vapor contents, T_x was the brightness temperature in x GHz, a_i, b_i, c_i, d_i denoted the coefficients of the functional relationship,they were also calculated using statistical regression approach (least square parameter estimation).

We can see, the equations from (4.19) to (4.22) are only different in the usage of frequencies. All of them include 22 GHz channel because the 22 GHz channel already explains most (98.7%) of the water vapor variance according to the regression results in (Schluessel and William, 1990). Additional measurements from other channels (19, 37, 86 GHz) are used to compensate for the perturbing effects of liquid water contents in the regression coefficients.

Finally Schluessel and William (1990) applied this estimated functional model to a set of real data measured by the satellite radiometer and compared the retrieval results of the brightness temperature with the values from radiosonde measurements from ships. At the end Schluessel and William (1990) concluded that the atmospheric water vapor contents could be calculated with an accuracy of between 0.145 and $0.17\mathrm{g/cm^2}$.

This statistical approach is further modified to provide path delay corrections for altimeter. We will see that the method used for altimetry is similar to the method described here, and the functional form in equation (4.19) is also preserved. Although atmospheric water vapor is a noise signal in altimetry and must be corrected through measurements by radiometer, the

byproduct of this range correction is a knowledge of total water vapor, since the measured brightness temperature by radiometer are at hand. So at the other hand equations (4.19) to (4.22) also provide us a method to acquire information about atmospheric water vapor.

4.4 Path delay for T/P

In the previous sections, the attention was focused on retrieval algorithms for total water contents. As total water vapor play an important role in the signal path delay, this retrieval concept is also used in altimetry for range corrections. Keihm et al. (1995) designed a pre-launch correction algorithms for T/P mission, a brief introduction of their work is given here.

Keihm et al. (1995) also used a statistical approach, just like the method used by Schluessel and William (1990) in section 4.3. However some modifications have been made in order to apply this method to altimetry:

1. Introduction of a path delay model,
2. Update of atmospheric and emissivity model,
3. Optimal frequencies analysis,
4. Two step statistical approach.

These modifications and results will be introduced in the rest part of this section.

4.4.1 Introduction of a path delay model

For range corrections in altimetry the focus must be transferred from water vapor contents to signal path delay. Keihm et al. (1995) modeled path delay due to total water vapor and total liquid water as follows:

$$PD_{\mathrm{v}} = 1.763 \cdot 10^{-3} \int_0^H \frac{\rho_{\mathrm{v}}}{T} dz, \tag{4.23}$$

$$PD_{\mathrm{l}} = 11.6 L_{\mathrm{z}}, \tag{4.24}$$

where PD_{v}, PD_{l} were path delay due to total water vapor and total liquid water, H was satellite height, ρ_{v} was the density of water vapor, T was the atmospheric temperature, L_{z} was the total liquid water in mm.

4.4.2 Update of atmospheric and emissivity model

An atmospheric model was also used to simulate the relevant parameter in equation (4.7), just like (Schluessel and William, 1990). But some modifications were taken due to updated experiments data. The sea surface emissivity was divided into three parts: roughness due to small scale and large scale waves, and roughness due to presence of foam. The last one was related to wind speed on the ocean.

4.4.3 Optimal frequencies analysis

A sensitivity analysis was added to find the optimum frequencies for range corrections used in T/P. The concern was that the interaction of atmosphere with electromagnetic waves also depended on the wave frequency, the optimal frequency was that at which the measurements are most sensitive to the change of atmospheric properties of interests, such as total water vapor. Keihm et al. (1995) tried to change relevant atmospheric state and observe how brightness temperature would response at frequency bands between 5-40 GHz.

At the beginning, Keihm et al. (1995) used a nominal model to simulate other surface and atmospheric parameters in equation (4.7) ($T_s = 295K$, wind speed $< 7m/s$, simplified atmospheric profile including temperature, pressure, density of water vapor and liquid water), so the measured brightness temperature could be simulated according to equation (4.7). Then about $10 - 20\%$ of the total expected variability in these parameter were inserted into the simulation, and the variation of simulated brightness temperature could be calculated. Base on the result illustrated in figure 4 in their paper, Keihm et al. (1995) made following consequences:

- the brightness temperature was not sensitive to the variation in surface pressure P_s and atmospheric temperature $T(z)$ from the nominal model at all frequencies between 5 and 40 GHz ($\delta T_b \approx 0$),
- a variation of 5 K in surface temperature T_s led to about 2 K at low frequencies but nearly zero at higher frequencies,
- if the wind speed changed $3m/s$, about 3 K variation in brightness temperature was observable at all frequencies ($5 - 40$ GHz),
- and the most important part, brightness temperature was dominant sensitive to variation in water vapor near the absorption line, about $20 - 25$ GHz,
- brightness temperature was only sensitive to variation in liquid water at higher frequencies.

Keihm et al. (1995) concluded that a frequency between $20 - 25$ GHz (most sensitive at 22.2 GHz) should be chosen to acquire a optimal sensibility to variation in water vapor, as in this spectral region variation in liquid water and wind speed also led to small but comparable variation in brightness temperature, so a higher frequency should be chosen to reflect the variation in brightness temperature due to liquid water, and a smaller frequency near absorption line should be chosen to reflect the variation in brightness temperature due to the wind speed. According to these considerations a combination 18, 22.2 and 37 GHz seemed to be optimal, but Keihm et al. (1995) replaced the 22.2 GHz with 21 GHz because at 21 GHz the measured brightness temperature was fast not sensitive to the height distribution of water vapor in the atmosphere at all, when the total water vapor is the object of interest, this is an important advantage. Finally the optimal frequencies were chosen as 18, 21, 37 GHz by Keihm et al. (1995).

4.4.4 Two step statistical approach

As other statistical approaches, measured brightness temperatures are usually simulated by physical models and radiosonde measurements, here Keihm et al. (1995) used a cloud model to calculate liquid density, used a four year radiosonde measurement data to calculate atmospheric temperature, pressure and water vapor density, the surface temperature and wind speed data were acquired from other missions (such as SSM/I, AVHRR/2), so that the path delays and brightness temperatures could be simulated using equation (4.23) and (4.7. Generally the relationship between these variables could be statistical estimated in a similar way presented in section 4.3.

The problem for statistical approaches is that the retrieval performance for anomalous weather conditions degrades due to large network interpolations errors. In order to reflect the effect of this anomalous weather conditions, especially rain and high-wind speed, Keihm et al. (1995) made some important modifications.

Given that the liquid water made only small contributions to path delay, a crude functional relationship was enough accurate for range correction due to liquid water. So Keihm et al. (1995) used site- and radiosonde-dependent average wind speed data to simulate brightness temperature, and estimated the coefficients in this function. At the same time a crude relationship between wind speed and simulated brightness temperature could also be estimated, so they obtained:

$$L = \sum l_i T_{\mathrm{b}}(\nu_i), \tag{4.25}$$

$$W = \sum w_i T_{\mathrm{b}}(\nu_i), \tag{4.26}$$

where L, W were total liquid water and wind speed, l_i, w_i were coefficients, $T_{\mathrm{b}}(\nu_i)$ was brightness temperature at frequency ν_i

In order to reflect the effect of anomalous weather conditions (high-wind speed), Keihm et al. (1995) simulated brightness temperature with different wind speeds: 0, 7, 14, 21, 28 m/s, for every different wind speed values, a corresponding brightness temperature data set was generated. The brightness temperature and path delay range could be associated in a functional relationship, whose coefficients could be estimated using simulated path delay and brightness temperature, so there were 5 groups of coefficients since 5 different wind speeds were used. These coefficients were called global coefficients.

Furthermore, within the data set for every wind speed, Keihm et al. (1995) stratified the simulated data set into 4 groups according to path delay ranges: $0-10, 10-20, 20-30$ and > 30. Coefficients are estimated for every range, because these coefficients were estimated for a given wind speed and path delay range ("locally" estimated), they were in general better than than global coefficients

This idea can be illustrated in the following figures:

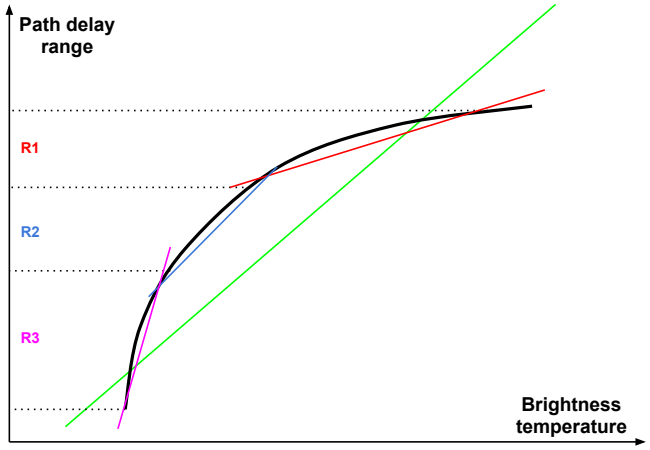


Figure 4.1: *Local and global estimation*

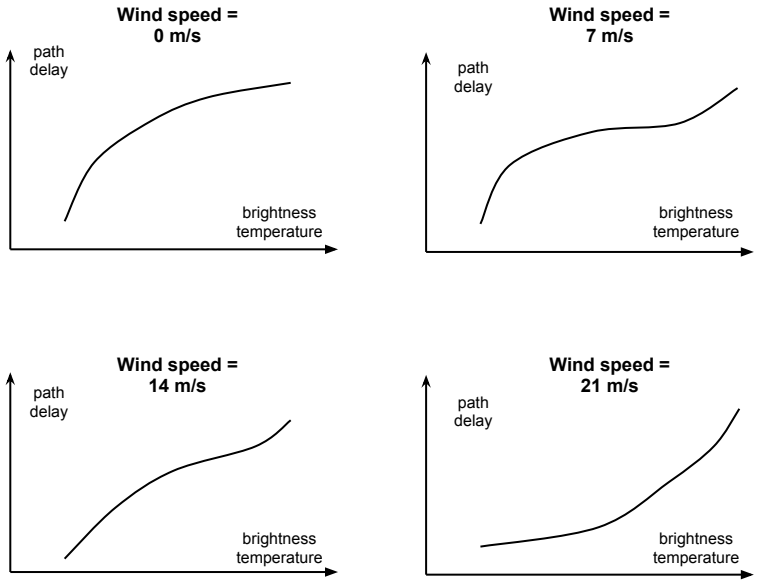


Figure 4.2: *Illustration: Relationship between brightness temperature and path delay depends on wind speed*

The high wind speed (anomalous weather conditions) can change the relationship between brightness temperature and path delay, this is also a problem for statistical approach as mentioned before. This difference caused by the wind speed is exaggeratedly illustrated in the figure (4.2). Their idea was to model every curve respectively, so they simulated brightness temperature with different wind speeds. For a given wind speed, global coefficients can be

estimated using all simulated data set at once, but errors are relative large, this situation is illustrated as a green straight line in the figure (4.1). In order to model the true black curve more accurately, the data set is divided into a few groups according to the path delay ranges, illustrated as $R1, R2, R3$ in the figure (4.1). Then coefficients for every range can be estimated using corresponding local data and the result is illustrated as red, blue and magenta straight lines in the figure. We can see that the combination of these local estimations is clearly better than just a global estimation.

At the time of range corrections, wind speed will at first approximately be calculated using equation (4.26) in order to distinguish which curve should be modeled. Then global coefficients are used to decide which path delay range. Finally the local coefficients are used to give an accurate range corrections. Contributions from liquid water is estimated using equation (4.24) and (4.25).

All retrieved coefficients were presented in a table (table II in their paper) and a interpolation technique could be further used to improve the performance. After a test and error analysis of this algorithm Keihm et al. (1995) stated that the overall measurements accuracy for wet range correction was about 1.2 cm.

We can see, this path delay correction algorithm is an application of radiative transfer theory and statistical retrieval approach introduced before, similar statistical methods are widely used in remote sensing to sound atmosphere. The textbook by Rees (2001) introduced a retrieval method for surface temperature using infrared signals, but he pointed out that this empirical retrieval method (for surface temperature retrieval) is only reliable provided that the emissivity of ocean surface is relative constant, so this method (surface temperature retrieval by statistical approach) can only be used for oceans because surface emissivity is relative constant there but not on land.

The remark by Rees (2001) can be considered as a explanation for this constraint. It means that a constant surface emissivity may be a prior condition for the usage of statistical retrieval approach. With this point of view we can review the work by (Keihm et al., 1995) in this section, their efforts of dividing data set by wind speed can be considered to be a strategy to use the dependence of statistical approach on surface emissivity, and their success also proves this dependence to some extent.

4.5 Update of radiometer in SWOT mission

Radiometer operating on frequencies between $18 - 37$ GHz have found successful application in oceanography. The mostly used statistical approach is an empirical method, which is designed and optimized for water surface on the ocean. It can not be used on land, or even $20 - 100$ km near coast. Dohan and Maximenko (2010) pointed out that the problem of the conventional radiometer in the application to coastal water surface was data contamination due to large footprint size of radiometer and inaccurate range correction due to atmospheric effect. In the literature, a few scientists are working on the update of radiometer for SWOT. The proposal concept is to augment the conventional radiometer to high frequencies. (Reising et al. 2011) and (Kangaslahti et al. 2010) discuss the design work of radiometer for SWOT in the future. According to their description, the radiometer used by SWOT will have following features:

- maintenance of the measurement principle of conventional radiometer operating on the frequencies between $18 - 37$ GHz,
- Augmentation of high frequencies centered at 92, 130 and 166 GHz for range correction over coastal areas,
- Augmentation of high frequency centerd at 183 GHz for range correction over land areas.

There will be three key technologies used on the radiometer:

- PIN-diode switch,
- high-excess noise radio noise source,
- single tri-frequency freed horn.

According to the (Reising et al., n.d.) and (Kangaslahti et al., 2010), these technologies will reduce the mass, power and volumn of the radiometer in SWOT, it will also be capable to perform internal calibration in the future.

Reising et al. (n.d.) and Kangaslahti et al. (2010) also provided an estimation of accuracy of the radiometer for SWOT: Over oceans and within 3 km from coast the accuracy of range correction is better than 1 cm; Over land the accuracy is about $1.5 - 3.0$ cm; Over large water bodies on land the accuracy is about 2.0 cm.

Reising et al. (n.d.) and Kangaslahti et al. (2010) said their design and test work was ongoing. At the moment the retrieval algorithms for land and coastal areas were not so clear. The concept of high frequencies for path delay retrieval with internal calibration had not been used for spaceborne platform befor. They only mentioned a so called Bayesian retriebal algorithm would be used for path delay retrieval over land and ocean.

Chapter 5

Applications of SWOT

This chapter will be devoted to applications of SWOT mission. The chapter will be divided into three sections for applications in oceanography, hydrology and others. Here some research themes and its actual knowledge will be briefly presented, shortages of some relevant measuring technique will also be mentioned, so that we will see the potential of SWOT mission in terms of applications in different areas.

5.1 Applications in oceanography

In the context of (Thomas et al., 2008), the scale on the order of kilometers is used as *submesoscale*, and scale of $10 - 100$ km is *mesoscale*. According to the description by (Fu et al., 2010a), the submesoscale is understood as scales of $1 - 100$ km, this range is larger than that given by (Thomas et al., 2008). Despite of this numeric difference, Thomas et al. (2008) characterized submesoscale processes as following: "An active flow field in the upper ocean generates localized regions, typically along filaments or outcropping isopycnals, within which the relative vertical vorticity equals or exceeds the planetary vorticity." Processes on this scale are important for turbulent transport and energy dissipation. However this scale has not fully been studied by researcher for lacking of measurement data. That is the main focus of SWOT mission in the future.

5.1.1 To measure submesoscale features of ocean

Dohan and Maximenko (2010) pointed out that global winds, Earth's rotation and the continental boundaries, these three factors set up the general ocean surface circulation. Although the long-term measurements of satellite altimetry has advanced our understanding of large scale circulation and heat storage of global oceans, large-scale surface motions are not completely described because of its complex interconnection of local currents, eddies, and turbulence.

An important surface feature are eddies, which can transport heat, salt, and biogeochemical properties such as nutrients and phytoplankton. Mesoscale eddies are expected to be responsible for about $20 - 30\%$ of vertical nutrient transport in the upper ocean (McGillicuddy et al., 2003). Despite of its importance they can not be resolved by a altimeter alone due to the poor spatial sampling of conventional altimeter (300 km equatorial spacing). Hence a combination of two or three altimeters was used to improve the spatial sampling resolution to observe the two-dimensional movement of mesoscale eddies. Based on these efforts the mesoscale eddies

with diameters larger than 100 km can be resolved by a combination of conventional altimeters. These researches enhance our knowledge on mesoscale features of ocean (Dohan and Maximenko, 2010). The ocean looks more like a sea of interacting eddies than a system of gyres and currents. Fu et al. (2010b) provided a comprehensive review of actual knowledge of ocean eddies from satellite altimetry and other remote sensing techniques. With the exceptions of a few localized regions, (mesoscale) eddies can be seen nearly everywhere in the world's ocean, and thousands of eddies can be identified globally at any given time, the merged data enable us to map the ocean eddies globally including its velocity, propagation pattern and transports of heat and salt. This argument can also be verified by a impressive snapshot of ocean eddies from merged data in two altimeter missions (Jason/Envisat), which was presented by Fu and Shum (2008). However this measurement method (multi-altimeter) is limited to the scale larger than 100 km. Ducet et al. (2000) presented global high-resolution maps of mesoscale variability using merged observations from multiple altimeters. The resolution of the merged data was only estimated as about 150 km in wavelength.

Besides eddies, most of the swift currents in the ocean have widths in the range of $10 - 100$ km, these have not been fully resolved (Fu et al., 2010b). Fu et al. (2010a) also pointed out that even with two altimeters the measurement resolution was not sufficient to resolve the details of the two-dimensional structure of ocean currents such as Gulf Stream and Kuroshio (on the order of 100 km).

In addition, at this submesoscale, there are many small eddies and high potential vorticity filaments which are shorter than 100 km and and about 1 km thick, they are stretched and pulled by the (large) eddies in ocean turbulence, these filaments are consistently seen in high-resolution SST images but have not been observed globally (Dohan and Maximenko, 2010). As a result of modeling studies for filaments, motions at these small scales are responsible for as much vertical transport as the large eddies (Klein and Lapeyre, 2009). Fronts are the interface between water masses of different densities, they also play an important role in the vertical transport in the upper ocean (Mahadevan and Tandon, 2006). Lapeyre and Klein (2006) pointed out that about 50% of the vertical transport of ocean biogeochemical properties was estimated to takes place at these submesoscale.

In the mission design, SWOT measurements will be averaged in 1 km $\times$ 1 km to achieve a height precision less than 2 cm (Fu et al., 2010a). So SWOT can provide oceanic submesoscale variability in any way, and that is also the most important contribution from SWOT.

5.1.2 Biogeochemical effect

As mentioned in the last section, the submesoscale features of ocean such as small eddies, filaments and fronts are important in vertical transport of ocean, this importance is mostly due to their large vertical velocity.

Submesoscale dynamics can generate vertical velocities of about 100 m/day which is typically an order of magnitude larger than those associated with the mesoscale. These intense vertical velocities can affect the lateral stirring of a phytoplankton bloom and enhance phytoplankton productivity through nutrient supply in filaments, they can also affect carbon export through the subduction of organic matter, and the air-sea transfer of carbon dioxide. At the same time, they generate lateral gradients at scales of kilometers that can enhance lateral mixing and create spatial heterogeneity in property distributions. Submesoscale processes are instrumental in

supplying nutrients for phytoplankton production, exchanging gases between the atmosphere and the ocean, and transferring heat, salt and momentum from the surface to the interior ocean (Thomas et al., 2008).

5.1.3 Flux of ocean energy

The energy budget of ocean is an important scientific question, at the mesoscale. A well established theory for energy cascade is quasi-geostrophic turbulence (Fu et al., 2010b), but at the submesoscale, the origin of energy is not well understood. The power spectrum of kinetic energy is often used to characterize the energy distribution in terms of length scales (Thomas et al., 2008). The change of slope of this power spectrum indicates an energy cascade between different scales. Thomas et al. (2008) reviewed recent results of studies by this approach. They also pointed out that it was still not well known how much energy of submesoscale flows was from atmospheric forcing, frontogenesis, and available potential energy release by instability.

An other important question is how energy will be transferred between different scales. Thomas et al. (2008) said, "a large part of the ocean's kinetic energy is located at meso and larger scales, at these scales, oceanic flow is already two-dimensional and in a state of hydrostatic and geostrophic balance from which it is difficult to extract energy." They pointed out that a scientific question had to be answered: "how energy is transferred from the mesoscale to the small-scale at which it can be dissipated through three-dimensional processes." This question is not well understood.

The oceanic processes at submesoscale can also affect the energy budget at large scale, but its knowledge is far from complete. For example, the cumulative effects of the submesoscale on larger scales have already been studied in terms of mixed layer restratification, but the effect of submesoscale on the potential vorticity budget at the large scale is not well known. It is supposed that the unbalanced submesoscale instabilities play an important role in the energy transition from mesoscale eddies to internal waves, but this argument also remains to be verified (Thomas et al., 2008).

5.1.4 Coastal measurements

The measuring principle of conventional altimetry is based on the wave analysis of echo signals reflected from ocean surface. The analyzing algorithm is optimized for the echos on the open ocean. But for the near coastal regions the echo signals are different from the echos on the open ocean, that causes error during wave analysis. An other problem for conventional altimetry is poor resolution, which leads to contamination of echo signals from ocean by noise from land surface near coastal regions. These drawbacks of conventional altimetry result in poor reliability of measurements near coastal regions. Observations by conventional altimeter are limited at distances of $25 - 100$ km offshore (Dohan and Maximenko, 2010). Several approaches have been invented to conquer this limitation, three of which are introduced by (Dohan and Maximenko, 2010). A method based on a combination of satellite data with in situ measurements and a statistical technique can even successfully calculate circulation up to 5 km from land. However, all these methods need regional treatment and depend on the regional studies (Dohan and Maximenko, 2010). SWOT can provide direct global measurement of water surface over coastal areas, it can also be used as complement to the regional studies.

Le Hénaff et al. (2008) investigated the potential qualitative improvement brought by wide-swath, interferometry-based ocean altimetry measurements in a coastal/shelf data assimilation system with respect to classical nadir altimeters. According to their study the entire shelf domain could be sampled in a shorter revisiting time in comparison with traditional altimeter. This allows for a better control of temporal changes in both sea level and velocities. They also pointed out that this improvement depended on the ocean state and the information provided by the interferometer overflights. The roll error of the platform should be taken into account and corrected if valuable observations were desired.

5.1.5 Validation for analytical methods of ocean dynamics

Isern-Fontanet et al. (2008) pointed out that an analytical theory called surface quasi-geostrophic theory (SQG) had been successfully used in extracting three-dimensional dynamics from surface fields. SQG has also been advanced as a appropriate model for submesoscale in terms of a forward energy cascade (Thomas et al., 2008). An application of this theory can be found in the work by Klein et al. (2009), they tried to diagnose the vertical velocity field from high-resolution Sea Surface Height (SSH) simulations using a surface Quasi-Geostrophy based (eSQG) method. Klein et al. (2009) were able to reconstruct low-frequency vertical velocities and horizontal motions within a range of scales between 20 km and 400 km from the surface down to 500 m, the necessary data were only a high-resolution SSH field and information on the large-scale vertical stratification. This work indicates a potential capability of SWOT to recover horizontal and vertical fluxes at submesoscale in the upper ocean. Measurements from SWOT then can also be used to validate its basic surface quasi-geostrophic theory (SQG) and advance its development.

5.1.6 Geostrophic currents

An application of conventional altimeter is to calculate the velocity of geostrophic currents. On scales of tens of kilometers and larger, horizontal motions are much larger than vertical motions, and the ocean is approximately hydrostatic. As a result of horizontal motions currents will subject to the Coriolis force, which drives currents that follow lines of constant pressure (Dohan and Maximenko, 2010). Therefore the velocity of geostrophic currents can be calculated from the pressure gradients, which are observed by altimeter as slope of ocean topography. Because conventional altimeters can only provide one-dimensional along-track data, so the velocity of geostrophic currents was calculated at crossover points by previous altimeters in order to determine a two-dimensional velocity field. With a wide swath SWOT can easily perform two-dimensional SSH measurement and make the calculation of the velocity of geostrophic currents easier.

5.2 Applications in hydrology

The scientific motivation of SWOT for hydrology is to answer three following question (Alsdorf et al., 2007):

1. What is the spatial and temporal variability in the world's terrestrial surface water storage and discharge? How can we predict these variations more accurately?

2. How much water is stored on a floodplain and subsequently exchanged with its main channel? How much carbon is potentially released from inundated areas?

3. What policy implications would freely available water storage data have for water management? Can health issues released to waterborne diseases be predicted through better mappings?

These three questions can be divided into two parts, one is the desire to special hydrological data, the other is the influence of this data on social and political areas. The second part is a logical implication of the first part. The most important is the first part, i.e. how to acquire the desired hydrological data (terrestrial water storage and discharge) globally. Two main approaches to estimate terrestrial water storage and discharge exist in situ methods and satellite based methods.

In situ method is traditionally used to measure water storage and discharge. The limitations of this method are (Fu et al., 2010a):

1. The distribution of global gauge stations are not homogeneous, most gauge stations are located in the industrialized countries. This results in a heterogeneous sampling of desired global hydrological data.

2. It is a point-wise method, the measurement values are only limited on some fixed points along a channel, the variability in the terrestrial water storage and discharge of the entire channel can not be accurately reflected only by data on these fixed points.

3. Some areas are not accessible and unmeasured, including many wetlands and floodplains.

4. Economic issue plays an important role, the number of gauge stations is declining worldwide.

5. The political boundaries hinder the data acquisitions and sharing. A new approach is needed to overcome these limitations of the traditional method (Fu et al., 2010a).

Before the proposal of SWOT, there have been a few methods from satellite based techniques. These methods can be divided into three groups. The first group is based on satellite infrared/visible band imaging technique which can provide direct area measurements of water bodies. The second group is based on conventional altimetry technique, which can provide direct elevation measurements of water surface. The third group is to measure the change of gravity field via GRACE. These three methods can not measure the water storage directly, but it can be calculated from area values, water surface height or change of gravity field respectively when the cross section of water bodies and other local information are combined with the measurements values.

Besides their dependence of the local information of water bodies, infrared imaging technique can not work in all-weather condition. Altimetry is a point-based method designed ideal for oceanographic applications, this technique can be used for inland water bodies but with degraded resolution. Gravity measurements via GRACE have a poor resolution of about 100 km, which is not suitable for application to most rivers and lakes. Alsdorf et al. (2007) provided

a good review of recent space-borne techniques, they noted that none of the existing methods alone could provide needed constraints on the global water cycle.

SWOT can make contributions to observe the variability of terrestrial water bodies. Besides the elevation measurements h, a 130 km wide swath enables SWOT to provide the slope or spatial gradient of water surface elevations $\partial h/\partial x$, a high temporal repeat rate (at least two times in 22 days globally, for mid-latitude areas even larger sampling rates) enables SWOT to measure the temporal change of water surface elevation $\partial h/\partial t$. These are the main products of SWOT for hydrology. In the literature there are some applications of these products to calculation of water storage and its temporal changes.

Hossain and Alsdorf (2011) analyzed the capability of SWOT measurement to monitor the water storage and its variability of a riverine delta called Bangladesh in South Asia. In their studies, the river discharge can be calculated by an empirical formula called Manning's Equation, which only requires $\partial h/\partial x$ and a few hydraulic parameters. Hossain and Alsdorf (2011) stated that the knowledge of water storage and storage discharge could help the local government to manage the water resources more effectively. Because the estimation of discharge by Manning's Equation also depends on a few hydraulic parameters, such as channel bathymetry, Manning's roughness coefficient, Durand et al. (2008) demonstrated a data assimilation methodology for estimating bathymetric depth and slope from water surface elevation measurements and the a hydrodynamic model. Furthermore, McCann et al. (2011) discusses a genetic approach to derive river bathymetry using only water surface elevations. Lee et al. (2010) evaluate storage change accuracy for the lakes in the Peace-Athabasca Delta, Northern Alaska and Western Siberia. An interesting conclusion made by them was that the storage change accuracy was insensitive to the orbital inclination or repeat periods, but was sensitive to lake shapes.

Besides the methods to calculate water storage from water surface elevations, as discussed above, the area extent of water bodies can also be determined from SWOT, Wu et al. (2011) used an airborne nadir KaRIN system to mimic the SWOT geometry and collect data over several diverse freshwater targets in the Van Hook Arm areas of North Dakota in USA, they found that the average return from water bodies was at least 20 dB greater than that from land, so the water area could be easily calculated using a threshold classification method. The look angle of this experiment was larger than in SWOT, so they said the contrast will be better for SWOT. This work suggests the capability of SWOT to measure the water area extent directly, so consequently the water discharge can then directly determined if the area extent and the elevation variability are combined together.

5.3 Byproduct of SWOT

SWOT is designed for two-dimensional measurements of SSH using InSAR technique. Derivations of the ideal situation have to be taken into account to achieve a high accuracy. This includes range corrections due to wet atmosphere and modulation of ocean surface by high-frequency tidal motions. These data are considered as noise and have to removed from raw data. On the other hand they also provide us data about wet vapor content and tidal motions as a byproduct of SWOT. So SWOT can also contribute to atmospheric and tidal studies.

Bibliography

Alsdorf, D., Rodríguez, E. and Lettenmaier, D. (2007), 'Measuring surface water from space', *REVIEWS OF GEOPHYSICS-RICHMOND VIRGINIA THEN WASHINGTON-* **45**(2).

Chang, A. and Wilheit, T. (1979), 'Remote sensing of atmospheric water vapor, liquid water, and wind speed at the ocean surface by passive microwave techniques from the nimbus 5 satellite', *Radio Science* **14**(5), 793–802.

Cumming, I. and Wong, F. (2005), 'Digital signal processing of synthetic aperture radar data: algorithms & implementation', *Recherche* **67**, 02.

Curlander, J. and McDonough, R. (1991), 'Synthetic aperture radar- systems and signal processing(book)', *New York: John Wiley & Sons, Inc, 1991.* .

Dohan, K. and Maximenko, N. (2010), 'Monitoring ocean currents with satellite sensors', *Oceanography* **23**(4), 94.

Ducet, N., Le Traon, P. and Reverdin, G. (2000), 'Global high-resolution mapping of ocean circulation from topex/poseidon and ers-1 and-2', *Journal of Geophysical Research* **105**(C8), 19–477.

Durand, M., Andreadis, K., Alsdorf, D., Lettenmaier, D., Moller, D. and Wilson, M. (2008), 'Estimation of bathymetric depth and slope from data assimilation of swath altimetry into a hydrodynamic model', *Geophysical Research Letters* **35**(20), L20401.

Esteban-Fernandez, D., Fu, L., Rodriguez, E., Brown, S. and Hodges, R. (n.d.), 'Ka-band sar interferometry studies for the swot mission development'.

Ferretti, A., Monti-Guarnieri, A., Prati, C., Rocca, F. and Massonet, D. (2007), *InSAR Principles-Guidelines for SAR Interferometry Processing and Interpretation*, Vol. 19.

Fu, L., Alsdorf, D., Rodriguez, E., Morrow, R., Mognard, N., Lambin, J., Vaze, P. and Lafon, T. (2010a), 'The surface water and ocean topography (swot) mission', *Proceedings of OceanObs* **9**.

Fu, L., Chelton, D., Le Traon, P. and Morrow, R. (2010b), 'Eddy dynamics from satellite altimetry', *Oceanography* **23**(4), 14–25.

Fu, L. and Shum, C.-K. (2008), 'Objectives and requirements of swot for observing the oceanic mesoscale variability', `http://bprc.osu.edu/water/Meetings_WATERHM/SWOT_Hydrology_Sept2008/Presentations/OSU-workshop-Fu.ppt`. [Online; accessed 25-April-2012].

Gentemann, C., Wentz, F., Brewer, M., Hilburn, K. and Smith, D. (2010), 'Passive microwave remote sensing of the ocean: An overview', *Oceanography from Space, Revisited, edited by V. Barale, JFR Gower and L. Alberotanza* pp. 13–33.

Grody, N. (1993), 'Remote sensing of the atmosphere from satellites using microwave radiometry', *Atmospheric remote sensing by microwave radiometry(A 95-14701 02-46), New York, NY, John Wiley & Sons, Inc., 1993*, pp. 259–334.

Hassen, R. (2001), 'Radar interferometry', *Netherlands: Kluwer Academic Publishers* .

Hossain, F. and Alsdorf, D. (2011), 'Understanding surface water flow and storage changes using satellites: Emerging opportunities for bangladesh', *Climate Change and Food Security in South Asia* pp. 57–67.

Isern-Fontanet, J., Lapeyre, G., Klein, P., Chapron, B. and Hecht, M. (2008), 'Three-dimensional reconstruction of oceanic mesoscale currents from surface information-art. no. c09005', *Journal of Geophysical Research-Oceans* **113**(C9).

Janssen, M. (1993), 'An introduction to the passive microwave remote sensing of atmospheres', *Atmospheric remote sensing by microwave radiometry(A 95-14701 02-46), New York, NY, John Wiley & Sons, Inc., 1993*, pp. 1–35.

Kangaslahti, P., Brown, S., Gaier, T., Dawson, D. and Harding, D. (2010), 'Radiometer testbed development for swot', *Earth Science* .

Keihm, S., Janssen, M. and Ruf, C. (1995), 'Topex/poseidon microwave radiometer (tmr). iii. wet troposphere range correction algorithm and pre-launch error budget', *Geoscience and Remote Sensing, IEEE Transactions on* **33**(1), 147–161.

Klein, L. and Swift, C. (1977), 'An improved model for the dielectric constant of sea water at microwave frequencies', *Antennas and Propagation, IEEE Transactions on* **25**(1), 104–111.

Klein, P., Isern-Fontanet, J., Lapeyre, G., Roullet, G., Danioux, E., Chapron, B., Le Gentil, S. and Sasaki, H. (2009), 'Diagnosis of vertical velocities in the upper ocean from high resolution sea surface height', *Geophysical Research Letters* **36**.

Klein, P. and Lapeyre, G. (2009), 'The oceanic vertical pump induced by mesoscale and submesoscale turbulence', *Annual Review of Marine Science* **1**, 351–375.

Lapeyre, G. and Klein, P. (2006), 'Impact of the small-scale elongated filaments on the oceanic vertical pump', *Journal of Marine Research* **64**(6), 835–851.

Le Hénaff, M., De Mey, P., Mourre, B. and Le Traon, P. (2008), 'Contribution of a wide-swath altimeter in a shelf seas assimilation system: Impact of the satellite roll errors', *Journal of Atmospheric and Oceanic Technology* **25**(11), 2133–2144.

Lee, H., Durand, M., Jung, H., Alsdorf, D., Shum, C. and Sheng, Y. (2010), 'Characterization of surface water storage changes in arctic lakes using simulated swot measurements', *International Journal of Remote Sensing* **31**(14), 3931–3953.

Liebe, H. (1985), 'An updated model for millimeter wave propagation in moist air', *Radio Science*, **20**, 1069–1089.

Mahadevan, A. and Tandon, A. (2006), 'An analysis of mechanisms for submesoscale vertical motion at ocean fronts', *Ocean Modelling* **14**(3), 241–256.

McCann, R., Andreadis, K., Alsdorf, D., Rodriquez, E. and Moller, D. (2011), A genetic approach to estimating river bed topography from swot observations, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International', IEEE, pp. 3039–3041.

McGillicuddy, D., Anderson, L., Doney, S. and Maltrud, M. (2003), 'Eddy-driven sources and sinks of nutrients in the upper ocean: Results from a 0.1 resolution model of the north atlantic', *Global Biogeochem. Cycles* **17**(2), 1035.

Parke, M., Stewart, R., Farless, D. and Cartwright, D. (1987), 'On the choice of orbits for an altimetric satellite to study ocean circulation and tides', *Journal of geophysical research* **92**(C11), 11693–11.

Rees, G. (2001), *Physical principles of remote sensing*, Vol. 1, Cambridge Univ Pr.

Reising, S., Kangaslahti, P., Brown, S., Dawson, D., Albers, D., Lee, A., Montes, O., Gaier, T., Hoppe, D. and Khayatian, B. (n.d.), 'Development of low-mass, low-power, high-frequency microwave radiometers with internal calibration to provide high-resolution wet-tropospheric path delay measurements for the swot mission'.

Rodriguez, E. (2008), 'Swot technology and expected performance', `http://bprc.osu.edu/water/Meetings_WATERHM/SWOT_Hydrology_Sept2008/Presentations/SWOT_OSU_Rodriguez.ppt`. [Online; accessed 17-Mai-2012].

Rodriguez, E., Pollard, B. and Martin, J. (1999), 'Wide-swath ocean altimetry using radar interferometry', *IEEE Transactions on Geoscience and Remote Sensing* .

Rosen, P., Hensley, S., Joughin, I., Li, F., Madsen, S., Rodriguez, E. and Goldstein, R. (2000), 'Synthetic aperture radar interferometry', *Proceedings of the IEEE* **88**(3), 333–382.

Schluessel, P. and William, J. (1990), 'Atmospheric water vapour over oceans from ssm/i measurements', *International Journal of Remote Sensing* **11**(5), 753–766.

Skolnik, M. (1970), 'Radar handbook'.

Sneeuw, N. (2006), 'Dynamic satellite geodesy', *Lecture notes, Geodätisches Institut, Universität Stuttgart* .

Soumekh, M. (1999), *Synthetic aperture radar signal processing*, Wiley New York.

Stammer, D. and Siegismund, F. (n.d.), 'Optimization of orbits for future eumetsat altimeter missions'.

Steven Nerem, R., Ryan, W., George, B. and James, C. (2008), 'Preliminary swot orbit design study', `http://bprc.osu.edu/water/Meetings_WATERHM/SubMesoscale_April2008/presentations-day2/Nerem.ppt`. [Online; accessed 01-May-2012].

Thomas, L., Tandon, A. and Mahadevan, A. (2008), 'Submesoscale processes and dynamics', *Ocean Modeling in an Eddying Regime, Geophys. Monogr. Ser* **177**, 17–38.

Wei, C., Leighton, H. and Rogers, R. (1989), 'Comparison of several radiometric methods of deducing path-integrated cloud liquid water', *Journal of Atmospheric and Oceanic Technology JAOTES* **6**(6).

Wentz, F. (1983), 'A model function for ocean microwave brightness temperatures', *Journal of Geophysical Research* **88**(C3), 1892–1908.

Wikipedia, the free encyclopedia (n.d.), `http://www.wikipedia.org`. [Online; accessed 17-Mai-2012].

Wu, X., Hensley, S., Rodriguez, E., Moller, D., Muellerschoen, R. and Michel, T. (2011), Near nadir ka-band sar interferometry: Swot airborne experiment, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International', IEEE, pp. 2681–2684.

Appendix A

Radiative transfer equation

The starting point of radiative transfer theory is to describe a radiation field in terms of the specific intensity I_ν. This is define as the instantaneous radiation power that flows at each point per unit area, per unit frequency interval and per unit solid angle as the radiation propagates from one point to another in the medium. The variation of the specific intensity can be described as a transfer equation:

$$\frac{dI_\nu}{ds} = -I_\nu \alpha + S, \tag{A.1}$$

where α is an absorption coefficient, S is a source term. On the right side hand of the transfer equation the first term describes the loss of energy and the second term describes the gain of energy. The equation is fulfilled in terms of total energy conservation.

S is the energy gain from an energy source, according to the Kirchoff's law:

$$S = \alpha B_\nu(T), \tag{A.2}$$

where $B_\nu(T)$ is the Planck function, which describes the amount of energy emitted by a black body in radiation of a certain wavelength.

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1}, \tag{A.3}$$

where h is Planck's constant, k is Boltzmann's constant, $c$is the speed of light, $\nu$is the frequency.

From the equation above, the transfer equation (A.1) can be solved as the radiation propagates along a given path from the boundary $s = s_0$ to the observer$s = 0$:

$$I_\nu(0) = I_\nu(s_0) e^{-\tau(s_0)} + \int_0^{s_0} B_\nu(T) e^{-\tau(s)} \alpha ds, \tag{A.4}$$

where τ is the optical depth or called opacity and defined as:

$$\tau(s) = \int_0^s \alpha(s') ds'. \tag{A.5}$$

Because the microwave frequency fulfills the *Rayleigh-Jeans limit*: $h \ll kT$, equation (A.3) can be approximated as:

$$B_\nu(T) \approx \frac{2kT}{\lambda^2}. \tag{A.6}$$

The Rayleigh-Jeans limit indicates a linear relationship of Planck function with physical temperature, so a new definition can be introduced according (A.6):

$$T_{\mathrm{b}}(\nu) = \frac{\lambda^2}{2k} I_\nu, \tag{A.7}$$

where T_{b} is called brightness temperature. This quantity is also the measurement by radiometer on TOPEX, Jason and SWOT, the water vapor can be calculated from this measured quantity, this will be showed later.

Using (A.7) the solution of the transfer equation(A.4) can be rewritten as:

$$T_{\mathrm{b}}(\nu) = T_{\mathrm{b0}}(\nu) e^{-\tau(s_0)} + \int_0^{s_0} \frac{T(s)}{R} e^{-\tau(s)} \alpha ds, \tag{A.8}$$

where T_{b0} is the background brightness temperature:

$$T_{\mathrm{b0}} = \frac{\lambda^2}{2k} I_\nu(s_0), \tag{A.9}$$

and the term R in the denominator inside the integral is:

$$R = \frac{2kT}{\lambda^2} \frac{1}{B_\nu(T)} = \frac{kT}{h\nu} (e^{h\nu/kT} - 1). \tag{A.10}$$

Because the Rayleigh-Jeans limit: $h \ll kT$, and the series expansion of e^x:

$$e^{h\nu/kT} - 1 = \frac{h\nu}{kT} + \left(\frac{h\nu}{kT}\right)^2 + \ldots \approx \frac{h\nu}{kT}, \tag{A.11}$$

so that the factor:

$$R \approx 1, \tag{A.12}$$

and then the solution of transfer equation under Rayleigh-Jeans approximation is:

$$T_{\mathrm{b}}(\nu) = T_{\mathrm{b0}}(\nu) e^{-\tau(s_0)} + \int_0^{s_0} T(s) \alpha e^{-\tau(s)} ds. \tag{A.13}$$

This equation has a clear physical meaning: T_{b0} is the background radiation. This radiation will be attenuated during the propagation through medium by a factor $e^{-\tau(s_0)}$, which depends on the distance between the background and observer. The second term at the right hand side is a summation of medium radiation energies which propagate from different distance away from observer and attenuated accordingly by a factor $e^{-\tau(s)}$.

This equation can be applied to the satellite case. We can assume the radiation propagates from the Earth's surface to the satellite height in the normal direction. Here the background is the Earth's surface and observer is the satellite radiometer, so the equation (4.1) can be rewritten as:

$$T_{\mathrm{b}}(\nu) = T_{\mathrm{b0}}(\nu) e^{-\tau(0)} + T_{\mathrm{u}}(\nu), \tag{A.14}$$

where T_{u} is the upwelling radiation term from the Earth's surface $z = 0$ to the satellite height $z = H$.

$$T_{\mathrm{u}}(\nu) = \int_0^H T(z) \alpha e^{-\tau(z)} dz. \tag{A.15}$$

The opacity $\tau(z)$ is rewritten for this case as:

$$\tau(z) = \int_z^H \alpha(z')dz'. \tag{A.16}$$

The background term T_{b0} can be divided into two components: radiation emitted from the Earth's surface T_e and downwelling radiation of atmosphere reflected from the Earth's surface and then back towards the satellite, which is denoted by T_r. So the (A.14) can then be rewritten as:

$$T_b(\nu) = T_u(\nu) + T_e(\nu)e^{-\tau(0)} + T_r(\nu)e^{-\tau(0)}. \tag{A.17}$$

Appendix B

Matlab codes for simulation of SWOT orbit

B.1 Aliased tidal signals

The following codes are for illustration of aliased tidal signal. The original frequency of hypothetical tide is modeled as 0.1 rad/s. The aliased sampling function is modeled as 0.005 rad/s. The Matlab function **fzero** is used to estimate the intersection point, function **fkt** describes the intersection point and is used as an input parameter for the function **fzero**.

Function: fkt

```
function f = fkt(x)
f = sin(0.005*x+pi/2)-sin(0.1*x);
```

Solution of intersection points

```
clear all
clc
x = zeros(1,25);
for i = 0:25
    x(i+1) =  fzero(@fkt,[5*pi + 20*pi*i,15*pi + 20*pi*i]);
end
```

Plot of aliased tidal signals

```
%Illustration of aliased tidal signals
t = 0:.02:480*pi;
figure;
subplot(2,1,1);
plot(t,sin(0.1*t))
hold on
plot(t,sin(0.005*t+pi/2),'r')
plot(x,sin(0.1*x),'ro','MarkerFaceColor',[.49 1 .63],'MarkerSize',6)
axis([0 300*pi -2 3])
xlabel('Time axis','FontSize',13)
ylabel('Tidal phase','FontSize',13)
legend('Tide signal','Aliased signal', 'Sampling phase','FontSize',13)
title('Illustration of a aliased tidal signal: different phase','FontSize',15)
```

```
box off
subplot(2,1,2);
plot(t,sin(0.1*t))
hold on
x2 = 5*pi+20*pi*(0:25);
plot(x2,sin(0.1*x2),'r');
plot(x2,sin(0.1*x2),'ro','MarkerFaceColor',[.49 1 .63],'MarkerSize',6)
axis([0 300*pi -2 3])
xlabel('Time axis','FontSize',13)
ylabel('Tidal phase','FontSize',13)
legend('Tide signal','Aliased signal', 'Sampling phase','FontSize',13)
title('Illustration of a aliased tidal signal: same phase','FontSize',15)
box off
```

B.2 Visualization of SWOT orbit

The following Matlab codes are for visualization of the SWOT orbit. This work has not taken nutation or precession into account. It presumes that the SWOT orbit is only a Kepler orbit. The equations in the codes are cited from the lecture script by Sneeuw (2006). The applied equations and the whole procedure are listed as follows:

1. mean angular speed:
$$n = \frac{\alpha}{\beta} \cdot \omega_E, \tag{B.1}$$

2. semi-major axis:
$$a = \sqrt[3]{\frac{GM}{n^2}}, \tag{B.2}$$

3. orbital period:
$$T = \frac{2\pi}{n}, \tag{B.3}$$

4. mean anomaly at epoch t:
$$M = M_0 + n \cdot \Delta t, \tag{B.4}$$

5. parameter transformation from Kepler elements to inertial Cartesian coordinates:
$$[\boldsymbol{r}_i, \boldsymbol{v}_i] = \text{Kepler2Kart}(a, e, I, \omega, \Omega, M), \tag{B.5}$$

6. transformation of inertial coordinates to earth fixed coordinates:
$$\boldsymbol{r}_e = R_3(\text{GAST}) \cdot \boldsymbol{r}_i, \tag{B.6}$$

7. transformation of earth fixed coordinates to geographic coordinates:
$$\lambda = \arctan(\frac{z}{\sqrt{x^2 + y^2}}), \quad \phi = \arctan(\frac{y}{x}). \tag{B.7}$$

The corresponding Matlab codes are as follows:

SWOT orbit simulation

```
%% Visualization of SWOT orbit
clear all
clc
rho = 180/pi;
% standard gravitational parameter
GM = 3.986005e14;
% Orbital eccentricity, inclination, argument of periapsis, longitude ...
% of the ascending node,
e = 0;
I = 78/rho;
omega = 268.3281/rho;
OMEGA = 48.1790/rho;
% Rotation speed of the Earth
rotv = 7292115e-11; % [rad/s]
% mean angular speed of satellite
n = 301/22*rotv;
% semi-major axis
a = (GM/n^2)^(1/3);
% period
T=2*pi/n;
% Mean anomaly at epoch t0
M0 = 91.6877/rho;
TT= M0 + [0:T/6000:320*T];
% Mean anomaly at epoch t
M = M0 + n*TT;
% Parameter transformation from Kepler elements to inertial coordinates
[ri,vi] = Kepler2Kart(a*ones(size(M)),e*ones(size(M)),I*ones(size(M)),...
    omega*ones(size(M)),OMEGA*ones(size(M)),M);
% Greenwich Apparent Sidereal Time
GAST = rotv*TT;
% Rotation matrix due to Earth rotation
R3 = rotationN(GAST,3);
% Cartesian coordinate in Earth-fixed coordinates:
re = multirot(R3,ri);
xe = re(:,1);
ye = re(:,2);
ze = re(:,3);
% geographic Coordinates
Breite = atan(ze./sqrt(xe.^2+ye.^2));
Laenge = atan2(ye,xe);
load coast.mat

%% 1 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]', 'FontSize',13)
title('2D Groundtrack on Worldmap: 1 Subcycles', 'FontSize',15)
hold on
for i=1:13
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
```

```
    hold on
end
axis tight
%% 2 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]','FontSize',13)
title('2D Groundtrack on Worldmap: 2 Subcycles', 'FontSize',15)
hold on
for i=1:13
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
    hold on
end
for i=14:27
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'g+','MarkerSize',1);
    hold on
end
axis tight
%% 3 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]','FontSize',13)
title('2D Groundtrack on Worldmap: 3 Subcycles', 'FontSize',15)
hold on
for i=1:13
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
    hold on
end
for i=14:27
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'g+','MarkerSize',1);
    hold on
end
for i=28:41
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'r+','MarkerSize',1);
    hold on
end
axis tight
%% 4 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]','FontSize',13)
title('2D Groundtrack on Worldmap: 4 Subcycles', 'FontSize',15)
hold on
for i=1:41
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
```

```
    hold on
end
for i=42:55
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'r+','MarkerSize',1);
    hold on
end
axis tight
%% 300 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]','FontSize',13)
title('2D Groundtrack on Worldmap: 300 Revolutions', 'FontSize',15)
hold on
for i=1:300
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
    hold on
end
axis tight
%% 301 Subcycles
figure;
plot(long,lat,'k','LineWidth',1.5),box off;
xlabel('longitude [degree]','FontSize',13)
ylabel('latitude [degree]','FontSize',13)
title('2D Groundtrack on Worldmap: 301 Revolutions', 'FontSize',15)
hold on
for i=1:301
    plot(Laenge((i-1)*6000+1:i*6000)*rho,Breite((i-1)*6000+1:i*6000)...
        *rho,'b+','MarkerSize',1);
    hold on
end
axis tight
```

The function **Kepler2Kart** transforms the Kepler elements to inertial coordinates, it uses following functions:

1. **inv_Keplergleichung**: calculate the eccentric anomaly,
2. **rotationN**: set up 3-D rotation matrices,
3. **multirot**: execute 3-D matrices multiplication.

These Matlab codes are listed respectively.

Function: Kepler2Kart

```
function [ri,vi]= Kepler2Kart(a,e,I,omega,OMEGA,M)

% Transformation of Kepler elements to Cartesian Position and Speed.
% a: semi-major axis [m]
% e: Eccentricity
```

```
% I: Inclination [Rad]
% omega: Perigee angle [Rad]
% OMEGA: Longitude of the ascending node. [Rad]
% M: mean Anomaly [Rad]
%
% ri: Position vector in Cartesian coordinate system [m]
% vi: Speed vector in Cartesian coordinate system [m/s]

% all vector in column vector transform:
a = a(:); e = e(:); I = I(:); M = M(:);
omega = omega(:); OMEGA = OMEGA(:);

GM = 3.986005e14;
E = inv_Keplergleichung(M,e);

% focal Coordinates:
rf = zeros(length(a),3);
rf(:,1)= a.*cos(E)-a.*e;
rf(:,2)= a.*sqrt(1-e.^2).*sin(E);

% mean rotation speed:
n = sqrt(GM./a.^3);
% focal speed:
vf = zeros(length(a),3);
vf(:,1)= -n.*a./(1-e.*cos(E)).*sin(E);
vf(:,2)=  n.*a./(1-e.*cos(E)).*sqrt(1-e.^2).*cos(E);

% in inertial coordinate system.
R3 = rotationN(-OMEGA,3);
R2 = rotationN(-I,1);
R1 = rotationN(-omega,3);
ri = multirot(R3,multirot(R2,multirot(R1,rf)));
vi = multirot(R3,multirot(R2,multirot(R1,vf)));
```

Function: inv_Keplergleichung

```
function E = inv_Keplergleichung(M,e)

% M: mean anomaly [Rad]
% e: eccentricity
% E: eccentric anomaly[Rad]

M = M(:); e = e(:);
E1 = M;
E2 = M + e.*sin(E1);
while max(abs(E1-E2))> 1e-11
    E1 = E2;
    E2 = M + e.*sin(E1);
end

E = E2;
```

Function: rotationN

```
function R = rotationN(alpha,rot_axis)

% alpha,rot_axis are vectors.
%
% Funktion rotationN: for every angle alpha(i) and its rotation axis
% rot_axis(i) will set up a rotation matrix R(:,:,i).

% Check dimension of alpha and rot_axis:
if length(alpha)==1
    alpha = alpha*ones(length(rot_axis),1);
elseif length(rot_axis)==1
    rot_axis = rot_axis*ones(length(alpha),1);
elseif length(alpha)~=length(rot_axis)
    cerr('Dimesion von rot_axis und alpha nicht uebereinstimmen!')
end

alpha = alpha(:);
rot_axis = rot_axis(:);

% N: how much rotation matrices must be calculated.
N = length(alpha);

% i: Index for every matrix.
i = (1:N)';

% R is a 3*3*N array,
% R(:,:,i) is the i-th rotation matrix
R(3,3,:) = 1:N;
R(1,1,:) = 1;   R(2,2,:) = 1;   R(3,3,:) = 1;

% They will be used later.
ai = mod(rot_axis,3)+1;
bi = mod(rot_axis+1,3)+1;

% Funktion "sub2ind"
% Transformation of Subscript(Vector) to Index(Integer).
% R(x,y,z)->R(Indexinteger)
R(sub2ind([3,3,N],ai,ai,i))=cos(alpha);
R(sub2ind([3,3,N],bi,bi,i))=cos(alpha);
R(sub2ind([3,3,N],ai,bi,i))=sin(alpha);
R(sub2ind([3,3,N],bi,ai,i))=-sin(alpha);
```

Function: multirot

```
function out_vec = multirot(rot_mat, rot_vec)

% function multirot defines a multiplication of a 3D array and
% a vector plane.
% rot_mat: 3*3*N array, rot_mat(:,:,i) is the i-th rotation matrix.
% rot_vec: N*3 array, rot_vec(i,:) is the i-th vector.
```

```
% out_vec: N*3 array, result.

N = size(rot_vec,1);
out_vec = zeros(3,N);
out_vec(1,:) = sum(squeeze(rot_mat(1,:,:)).*rot_vec');
out_vec(2,:) = sum(squeeze(rot_mat(2,:,:)).*rot_vec');
out_vec(3,:) = sum(squeeze(rot_mat(3,:,:)).*rot_vec');

out_vec = out_vec';
```