MARSIS Radar: the Contrast Method Algorithm

OLD CATALOGUE:
INAF/IAPS-2014-02 / ISSUE 1 / REVISION 0

Issue 1, Rev 0

PREPARED by : Marco Cartacci[1], Andrea Cicchetti[1], Raffaella Noschese[1]

CHECKED by : Roberto Orosei[2]

APPROVED by : Roberto Orosei[2]

[1]INAF-IAPS Via Fosso del Cavaliere 100, 00133, Rome, Italy
[2]INAF-IRA Via Piero Gobetti, 101, 40129 Bologna, Italy

1. Introduction

It is well know that Mars does not possess an appreciable global magnetic field, by contrast to Earth, which has a strong geomagnetic field of core origin. In these conditions, the solar wind can directly interact with the Martian ionosphere and induce modifications of its local properties, producing a clear distinction between the day side and the night side ionosphere behaviors . This is mainly due to the fact that the day side is directly hit by solar EUV photons which ionize atmospheric neutrals.

As a consequence, the use of a radar sounder to analyse the surface and subsurface of Mars, through the ionosphere, must deal with serious constraints, depending on the operative frequencies adopted.

In this document, we describe the Contrast Method algorithm used for the on-board and the on-ground processing of the MARSIS radar, in order to compensate the distortions introduced by the Mars ionosphere on the data collected by MARSIS in its subsurface operation mode. We show the effects obtained with the Contrast Method in different environmental. The result is that, the presence of the Contrast Method in the processing pipeline has allowed the collection of relevant scientific data not only during the night side, but also during the day side, producing a substantial improvement of the data quality and the Mars coverage.

2. The MARSIS instrument

The Mars Advanced Radar for Subsurface and Ionosphere Sounding (MARSIS) (Picardi et al. 2005), carried by ESA's Mars Express spacecraft, is a nadir-looking pulse limited radar sounder, which uses synthetic aperture (SAR) techniques. MARSIS was developed by the University of

Rome "La Sapienza", Italy, in partnership with NASA's Jet Propulsion Laboratory in Pasadena, California. The main task of the MARSIS mission is to find evidences of the presence of water, both liquid and solid, on Mars, with the secondary objective of characterizing the structure of the Martian ionosphere. In order to achieve these goals, MARSIS has two operation modes: the SS (Sub-Surface) Mode and the AIS (Active Ionosphere Sounding) Mode.

In its SS (Sub-Surface) mode, MARSIS transmits "chirps" (linear FM), i.e. wave packets of duration T = 250 μsec which are linearly modulated in frequency over a bandwidth B = 1MHz, centred at 1.8 MHz, 3 MHz, 4 MHz or 5 MHz, alternating the transmission at two different frequencies, from a 40-m dipole antenna with a Pulse Repetition Frequency (PRF) of 127 Hz. MARSIS is optimized for deep penetration, so the frequencies are chosen according to the predicted SEA (Sun Elevation Angle, i.e. the angle between the direction of the geometric center of the Sun's apparent disk and the horizon), in order to satisfy the constraint to have the chirp frequency always higher than the local plasma frequency.

The radar vertical resolution after the range compression (which is defined as the convolution between the received chirp signal and a reference function representing the emitted chirp), and after Hanning windowing that is applied to reduce the amplitude of the pulse side-lobes, is approximately 210 m in the free-space. In the subsurface, the resolution is improved by a factor equal to the square root of the soil permittivity, assuming values of 50-100 m.

The SAR processing is designed in order to obtain synthetic apertures (called frames) adjacent to each other, with a ground resolution of 5.5-10 km along the track and of 17-30 km across the track, where lower and higher resolutions pertain to higher and lower S/C altitudes, respectively. The receiving window duration is 350 μs and the sampling frequency is f_s = 1.4MHz, so that each frame contains 490 samples that increase to 512 after zero padding and FFT processing.

During each synthetic aperture, MARSIS alternates two frequencies at PRF steps, so as to obtain more information about subsurface characteristics and to increase the probability that at least one of them propagates above the plasma frequency: the higher frequency (F01) is emitted before the lower one (F02). One additional feature of MARSIS is that it is equipped with a tracking loop that allows the radar to keep echoes within the receiving window in order to follow the surface profile, regardless of the presence of any additional ionospheric delay.

With the exception of ice, the penetration depth of radar signals in the subsurface is approximately proportional to their wavelength, as a conseguence MARSIS operates at the lowest possible frequencies capable of propagating through the martian ionosphere, i.e. just above the local plasma frequency, f_p. As the electron density is known to be definitely lower in the night side, this constraint implies that the MARSIS subsurface sounder is best utilized for negative values of the SEA.

Actually, along an orbit, the MARSIS radar performs a large number of synthetic apertures. Each synthetic aperture consists of about 200 radar pulses (this value depends on different parameters as signal frequency, altitude, tangential velocity etc.), transmitted in a segment of orbit approximately 5,5 km long (minimum value for a single synthetic aperture). Through unfocused SAR processing (azimuth and range compression), the measurements (the echo received for each pulse) are reduced to a single radar trace (frame) which represents the power backscattered from the surface and subsurface discontinuities versus the two-way travel time. A continuous sequence of frames produces the so called “radargram” where the X-axis represents the satellite orbital direction and the Y-axis the two-way travel time, while the power of the signal is coded by the grey scale color map.

3. The effects of the martian ionosphere on MARSIS signal propagation

The presence of the martian ionosphere produces a variation of the refraction index respect to the vacuum. As a consequence, an electromagnetic wave of frequency f propagating in the ionosphere is characterized by the following refraction index

$$n(z) = \sqrt{1 - \frac{f_p^2(z)}{f^2 - jf\nu}} \cong \sqrt{1 - \frac{f_p^2(z)}{f^2}} \quad (1)$$

where f_p is the plasma frequency, ν the electron-neutral collision frequency and z is the altitude above ground. Considering a typical MARSIS operation frequency (i.e. in the 1.3-5.5 MHz range), the imaginary term in the denominator of Eq. (1) can be neglected, because $\nu \sim 10 - 60$ kHz. The plasma frequency, in Hz, can be written as

$$f_p(z) = 8.98\sqrt{N_e(z)}, \quad (2)$$

where N_e is the electron density in m^{-3}. The maximum value of f_p obviously corresponds to the maximum value of the electron density N_{emax}.

Actually, according to Eq. (1) all frequencies lower than f_p will be reflected, while if the radar signal has a wide band, the propagation speed is not constant through the band itself and a frequency dependent phase shift arises. In details, frequencies higher than f_p will be attenuated, delayed by an average delay (group delay) in signal travel time and dispersed depending on the electron density values encountered along the path.

The phase shift induced by the ionosphere in a radar signal of frequency f can be written as

$$\Delta\phi(f) = \frac{4\pi}{c} f \int_0^L [n(z) - 1]\, dz = \frac{4\pi}{c} f \int_0^L \left[\sqrt{1 - \left(\frac{f_p(z)}{f} \right)^2} - 1 \right] dz, \quad (3)$$

where L is the ionosphere thickness and c is the speed of light in vacuum.

If f_0 is the central frequency of the radar signal band, we can perform a Taylor expansion of the integrand of Eq. (3) and then integrate each term of the expansion, so as to obtain

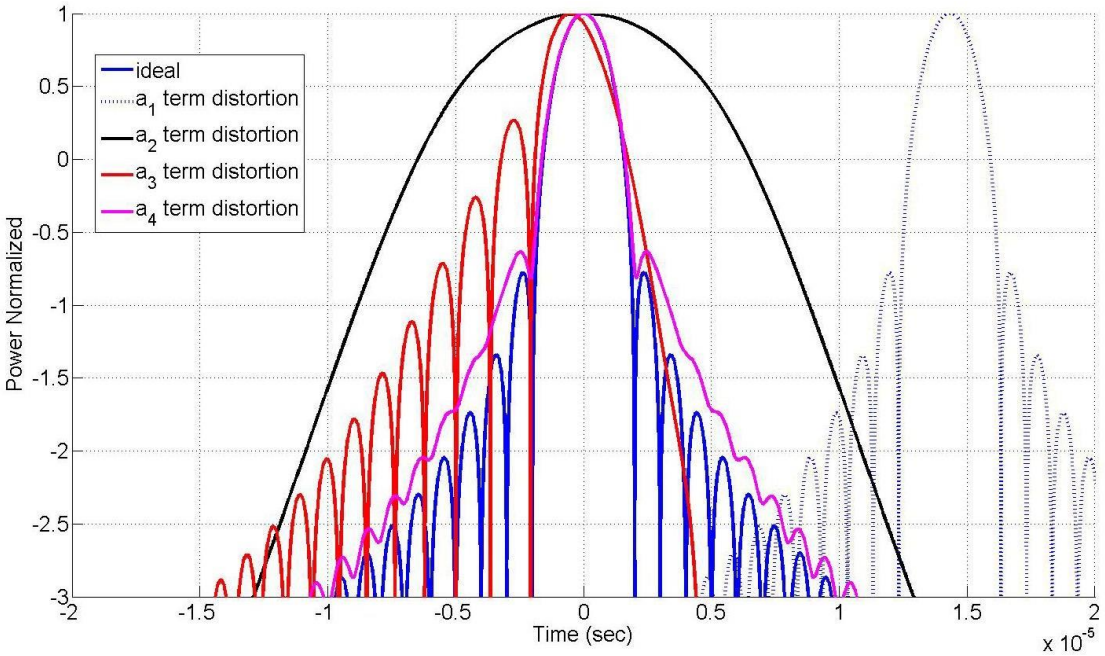


Fig. 1. Simulation of the effect of phase distortion on an ideal radar signal. Blue line: convolution of the ideal reflected "chirp"; blue, black, red and purple lines: convolutions pertaining to the a_1, a_2, a_3, and a_4 expansion terms (see Eqs. 16 - assuming that f_p = 1 MHz and τ_0 = 533 μs). All lines have been normalized to their peak values.

$$\Delta\phi(f) \cong a_0 + a_1(f-f_0) + a_2(f-f_0)^2 + a_3(f-f_0)^3 + a_4(f-f_0)^4 + \ldots, \quad (4)$$

where:

$$a_0 = \frac{4\pi}{c}\int_0^L \left(\sqrt{f_0^2 - f_p^2} - f_0\right) dz \quad [rad]$$

$$a_1 = \frac{4\pi}{c}\int_0^L \left(\frac{f_0}{\sqrt{f_0^2 - f_p^2}} - 1\right) dz \quad [rad/Hz]$$

$$a_2 = -\frac{4\pi}{c}\int_0^L \left(\frac{f_p^2}{2(f_0^2 - f_p^2)^{\frac{3}{2}}}\right) dz \quad [rad/Hz^2] \quad (5)$$

$$a_3 = \frac{4\pi}{c}\int_0^L \left(\frac{f_0 f_p^2}{2(f_0^2 - f_p^2)^{\frac{5}{2}}}\right) dz \quad [rad/Hz^3]$$

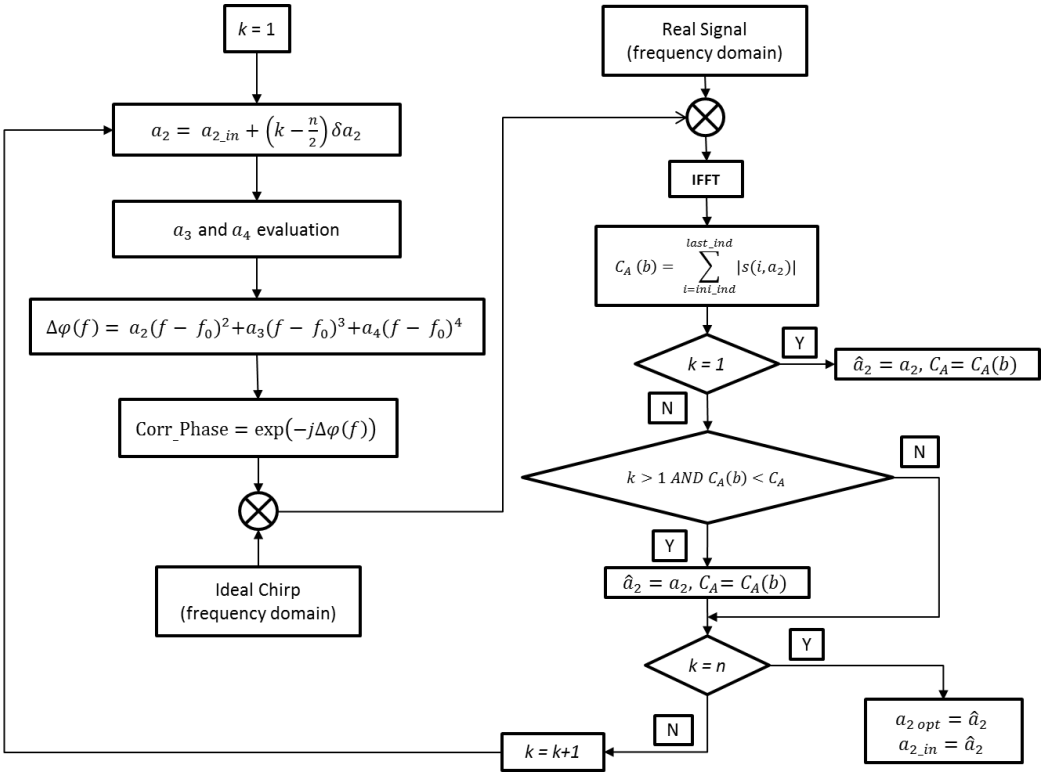
$$a_4 = \frac{4\pi}{c}\int_0^L \left(\frac{4 f_0^2 f_p^2 + f_p^4}{8(f_0^2 - f_p^2)^{\frac{7}{2}}}\right) dz \quad [rad/Hz^4]$$

The effect of the $a_0, \ldots, a_4$ expansion coefficients on the MARSIS SS performance can be briefly described by recalling that the SS data are processed through the range compression. In the ideal case, after range compression and Hanning weighting, the theoretical main lobe width, at 3 dB under the peak, should be about 1.5 μsec wide, while the difference in power between the main lobe peak and the first side lobe peak should be 32 dB. These parameters characterize the radar range resolution, that is the ability to reveal objects close to one another, and the radar dynamic range, which affects the capability to detect subsurface echoes.

Fig. 1 displays the range compression of the ideal reflected signal (i.e. in free space - blue line) and the range compressions pertaining to the a_1, a_2, a_3, and a_4 expansion terms (blue dotted, black, red and purple lines, respectively), all evaluated through the simplified expressions given in Eqs. (16)

of the following section, assuming that f_p = 1 MHz and τ_0 = 533 μs. All lines have been normalized to their peak values. We notice that the a_1 term only introduces a time shift (group delay), while the higher terms yield phase distortions. In particular, the a_4 term smears out the side lobes, while the a_3 term enhances the lobes preceding the main lobe and reduces those following it. The most relevant effect is that due to the a_2 term, which seriously affects the received chirp slope.

In conclusion, we see that the ionosphere can severely degrade the data quality, i.e. increase the side lobe levels, distort the waveform shape, and worsen the signal to noise ratio and range resolution. Moreover, the MARSIS signal is very vulnerable to ionosphere effects especially in those areas where the ionosphere and the magnetic field effects are combined together, because in these areas distortions are larger than usual. Obviously all the effects will increase steadily passing from the night side to the day side.



4. The Contrast Method

Where n is the number of iterations, $\hat{a}_{2start}$ is the starting value and δa_2 the iteration step. $\hat{a}_{2start}$ is set to zero for the first frame of a given orbit, while for all the following frames it is set equal to the best value of $\hat{a}_2$ estimated in the preceding frame.

In order to estimate an upper value for δa_2, let us define the effect of the ionosphere distortion as:

$$\Phi_1(f) + \Phi_D(f) = \Phi_1'(f) \qquad (8)$$

Where $\Phi_1(f)$, $\Phi_D(f)$ and $\Phi'_1(f)$ are the square-law phase terms of the transmitted signal, of the ionosphere distortion and of the received signal. Considering Eq. (6-20) from Cook and Bernfeld (1967) and Eq. (5), we can write:

$$\Phi_1(f) = \frac{4\pi^2(f - f_0)^2}{2\mu}$$

$$\Phi_D(f) = a_2(f - f_0)^2 \quad (9)$$

$$\Phi_1'(f) = \frac{4\pi^2(f - f_0)^2}{2\mu'}$$

μ is the slope of the transmitted chirp signal and can be expressed as

$$\mu = \frac{2\pi B}{T} \quad (10)$$

where B = 1 MHz and T = 0.25 ms are the chirp bandwidth and duration, respectively, while μ' is the slope of the received chirp.

Eqs. (8) and (9) yield

where the mismatching factor between transmitted and reflected chirp is

$$\gamma' = \frac{\mu - \mu'}{\mu'}$$

We can define the accuracy of the phase correction as:

$$\Delta a_2 = a_2 - \hat{a}_2 = \frac{\pi\gamma' T}{B} - \frac{\pi\hat{\gamma}T}{B} = \frac{\pi T}{B}(\gamma' - \hat{\gamma}) = \frac{\pi T}{B}\gamma \quad (13)$$

Where $\hat{a}_2$ and $\hat{\gamma}$ are the estimated correction term and the estimated mismatching factor, respectively, while γ is the residual mismatching factor. From Fig. 6.27, pag. 156 from Cook C.E., Bernfeld M., 1967, we can assume, as a worst case, that:

$$\gamma \leq \frac{2}{BT} \quad (14)$$

In conclusion, Eqs. (13) and (14) yield

$$\Delta a_2 \leq \frac{2\pi}{B^2} = 6.28 \cdot 10^{-12} \quad rad/Hz^2 \quad (15)$$

In practice, in the Eg. (7) iteration, it is desirable to use a step smaller than Δa_2.

Usually, $\delta a_2 = 0.1\ \Delta a_2$. In order to determine the $\hat{a}_3$ and $\hat{a}_4$ terms to be used in the iteration defined in Eq. (6), we first simplify Eq. (5) by making use of a model ionosphere characterized by a constant plasma frequency $f_{p,max}$ and an equivalent slab thickness $L_{Eq.}$ This assumption allows to move the integrand out of the integral, which becomes trivially equal to L_{eq}, so that we obtain (12)

$$a_1 = 2\pi\tau_0\left(\frac{f_0}{\sqrt{f_0^2 - f_{p\max}^2}} - 1\right) \quad [rad/Hz]$$

$$a_2 = -2\pi\tau_0\left(\frac{f_{p\max}^2}{2(f_0^2 - f_{p\max}^2)^{\frac{3}{2}}}\right) \quad [rad/Hz^2] \qquad (16)$$

$$a_3 = 2\pi\tau_0\left(\frac{f_0 f_{p\max}^2}{2(f_0^2 - f_{p\max}^2)^{\frac{5}{2}}}\right) \quad [rad/Hz^3]$$

$$a_4 = -2\pi\tau_0\left(\frac{4f_0^{\,2} f_{p\max}^2 + f_{p\max}^4}{8\left(f_0^2 - f_{p\max}^2\right)^{\frac{7}{2}}}\right) \quad \left[rad/Hz^4\right]$$

having defined $\tau_0 = 2L_{eq}/c$.

Assuming that $(f_{pmax}/f_0)^2 << 1$, from Eq. (16) we easily find that

$$\hat{a}_3 \cong -\frac{\hat{a}_2}{f_o}\left(1 - \frac{\hat{a}_2 f_o}{\pi\tau_o}\right) \qquad (17)$$

$$\hat{a}_4 \cong \left(\frac{\hat{a}_2}{f_0^{\,2}}\right)\left(1 - \frac{\hat{a}_2\ f_0}{0.5\pi\ \tau_0}\right)$$

Actually, the compensation term $\Delta\varphi$ that produces the range compressed signal with the best energy concentration in a defined time interval of the receiving window, is selected to perform the final range compression. In practice, the CM provides $\hat{a}_{3,}$ $\hat{a}_3$ and $\hat{a}_4$ as best estimates of the coefficients of the expansion defined by Eq. (4).

The Contrast Method is applied to all synthetic apertures (frames) collected by MARSIS and for each frequency. Fig. 2 shows how the CM improves the quality of the range compressed data for a given frame: the black line shows the signal after range compression without correcting the phase through the CM, while the blue line shows it after the CM optimization procedure has been applied. It is evident that the CM yields a higher peak power, a better signal to noise ratio and a reduction of the main lobe width, leading to a better range resolution; this allows to separate the surface and subsurface echoes that without correction would be merged together.

It is worth noting that the CM does not compensate the time delay introduced by the ionosphere; so in the signal radargrams shown in the figures introduced in the following sections, the time delay has been corrected by other means.

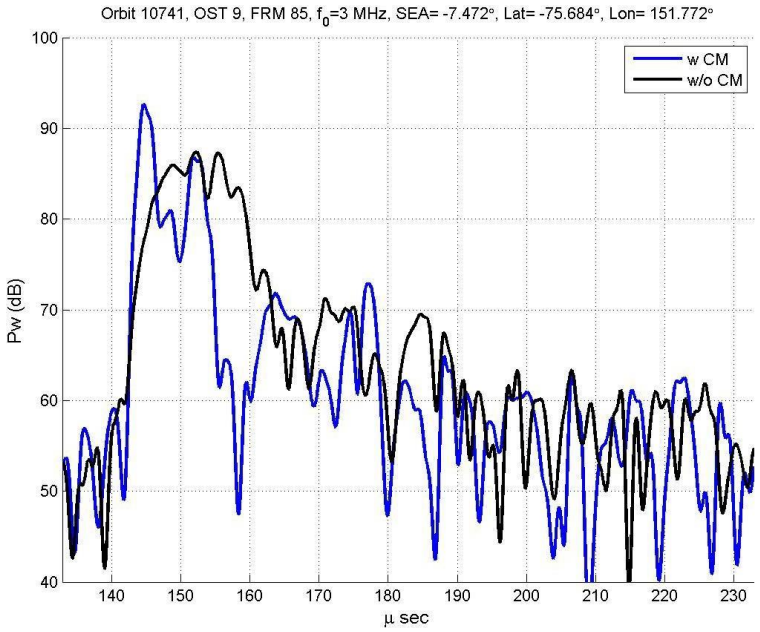


Fig. 2. Reflected power as a function of time during frame 84 of orbit 10741. The black line shows the uncorrected received signal; the blue line shows the reflected power after correction through the Contrast Method.

5. The Contrast Method Performances Analysis

The example shown in Fig. 2 introduced at the end of the preceding section clearly suggest that the CM is fundamental to obtain data scientifically useful. In Figs. 3, 4 there are two radargrams collected in two separate segment of orbit 10741 and characterized by different surface and SEA behavior. In Fig. 3, the frames has been collected during the night side with a $-14° \leq SEA \leq -7°$ and a frequency of 3 MHz; the panel a) shows the signals after the ionosphere distortion correction through the Contrast Method while in the panel b) the ionosphere distortion has not been compensated.We notice that after the frames number 40 the signal in the panel b) is blurred and the surface and the subsurface layers are merged together while in the panel a) are both clearly visible. This effect is due to the fact that, as anticipated in Fig. 2, the signals lobes are broaden by the ionosphere distortion.

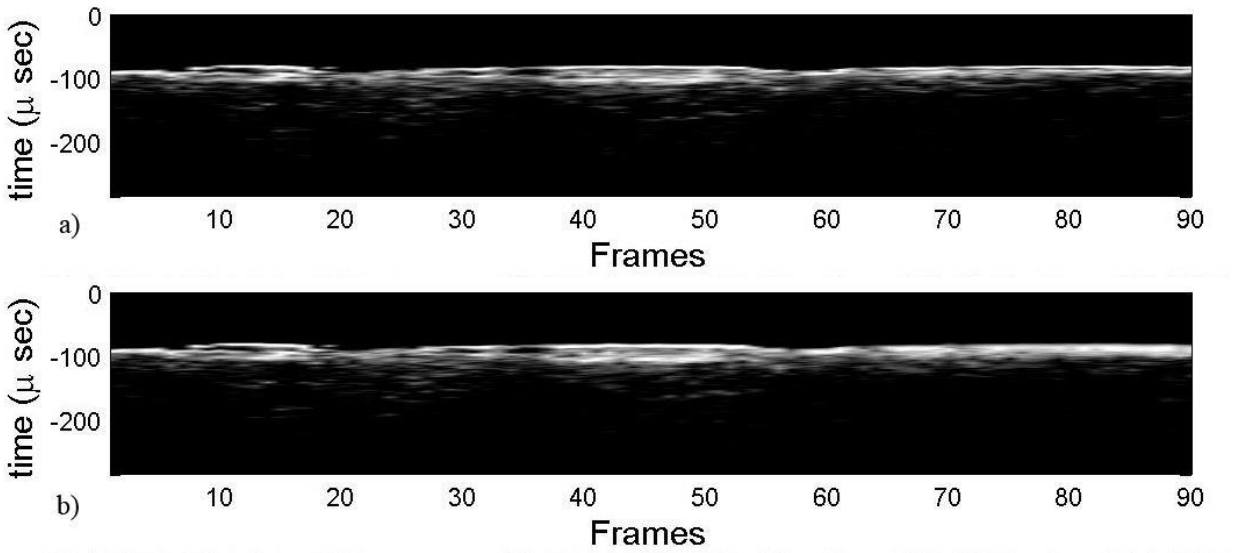


Fig. 3. The radargrams represented, pertain to the orbit 10741, the frames has been collected during the night side with a -14°≤SEA≤-7° and a frequency of 3 MHz; the panel a) shows the signals after the ionosphere distortion correction through the Contrast Method while in the panel b) the ionosphere distortion has not been compensated.

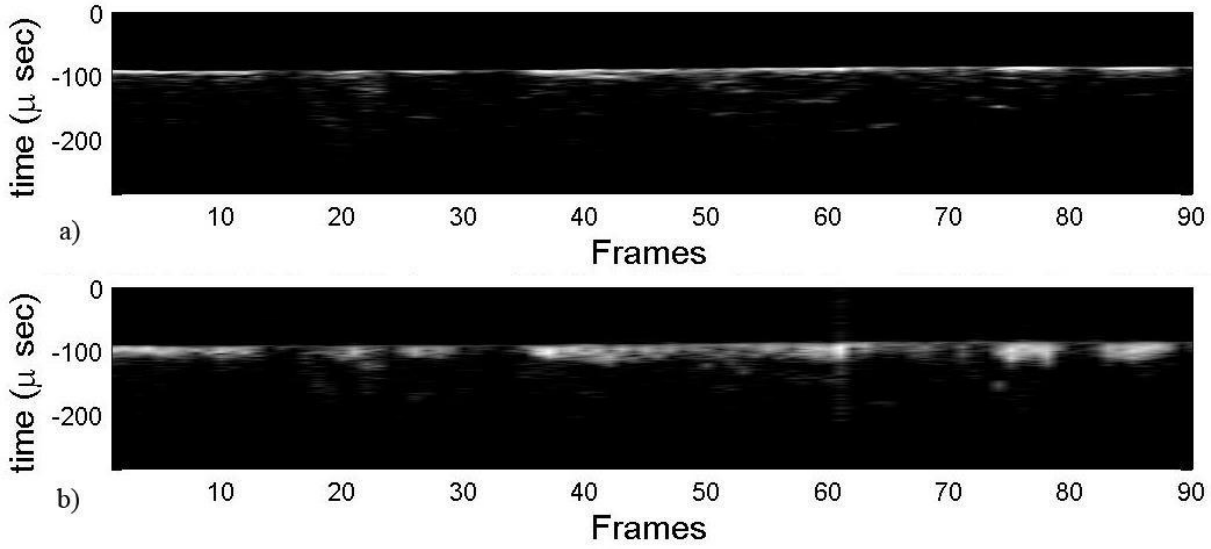


Fig. 4. The radargrams represented, pertain to the orbit 10741, the frames has been collected during the night side with a 0°≤SEA≤7° and a frequency of 4 MHz; the panel a) shows the signals after the ionosphere distortion correction through the Contrast Method while in the panel b) the ionosphere distortion has not been compensated.

In Fig. 4, the data has been collected during the day side with a 0°≤SEA≤7° and a frequency of 4 MHz, in this situation the effect of the ionosphere distortion is much higher, in fact the surface echoes in the panel b) are at least four time wider than the same echoes represented in panel a). The further step was to quantify the improvement obtained by using the CM. To this aim, Figs. 5 and 6 show the SNR behavior for the same frames shown in Figs. 3 and 4 respectively.

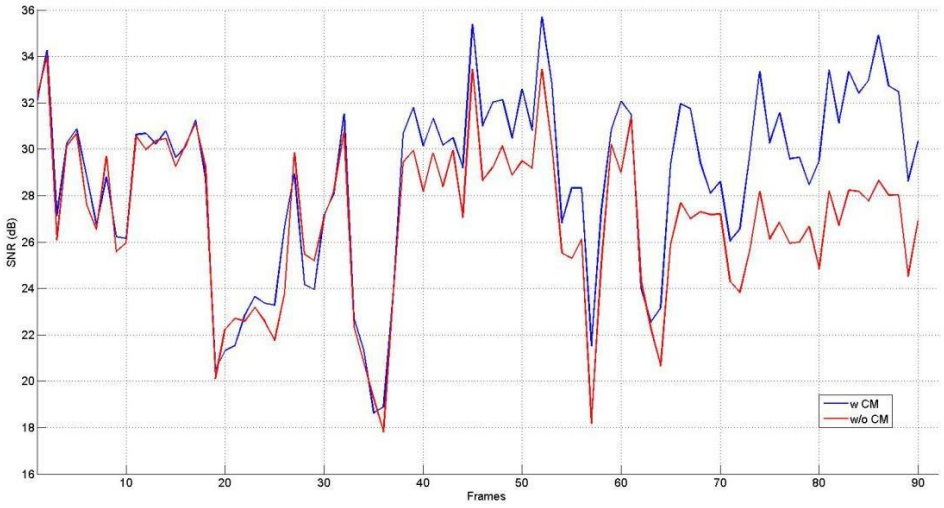


Fig. 5. SNR behavior, for 3 MHz, connected to the radargrams of Fig. 3. The blue and the red lines show the SNR obtained with and without the CM respectively.

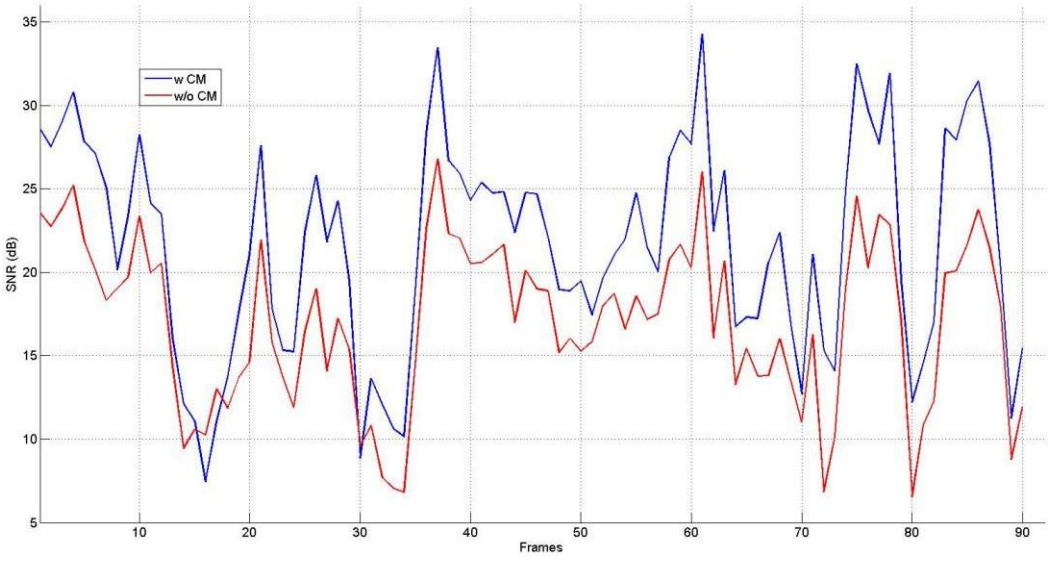
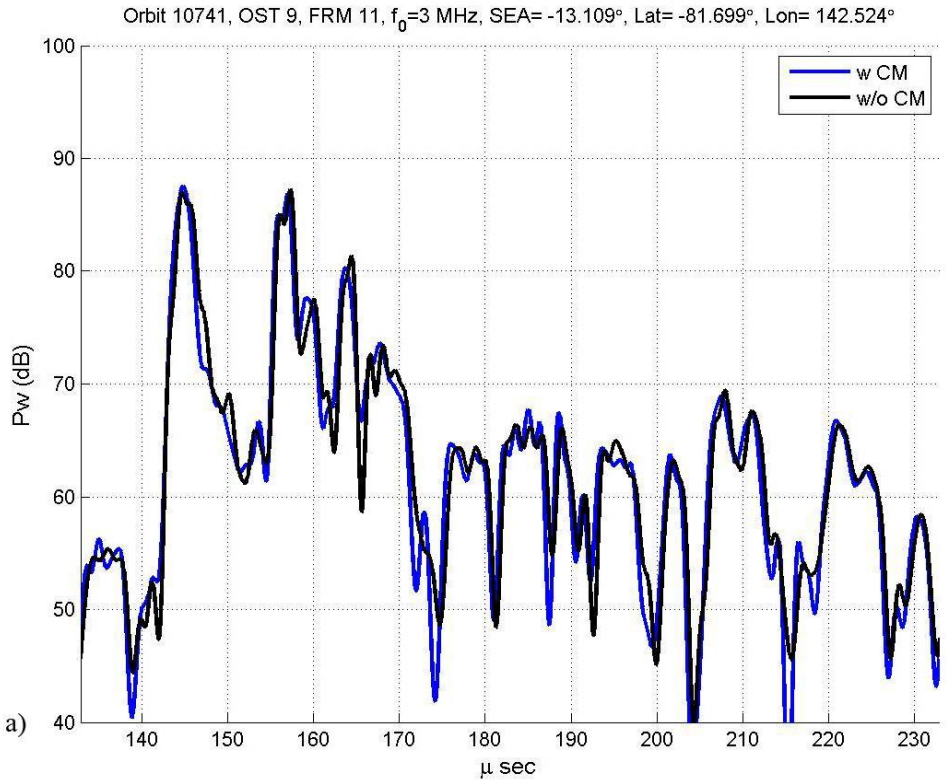


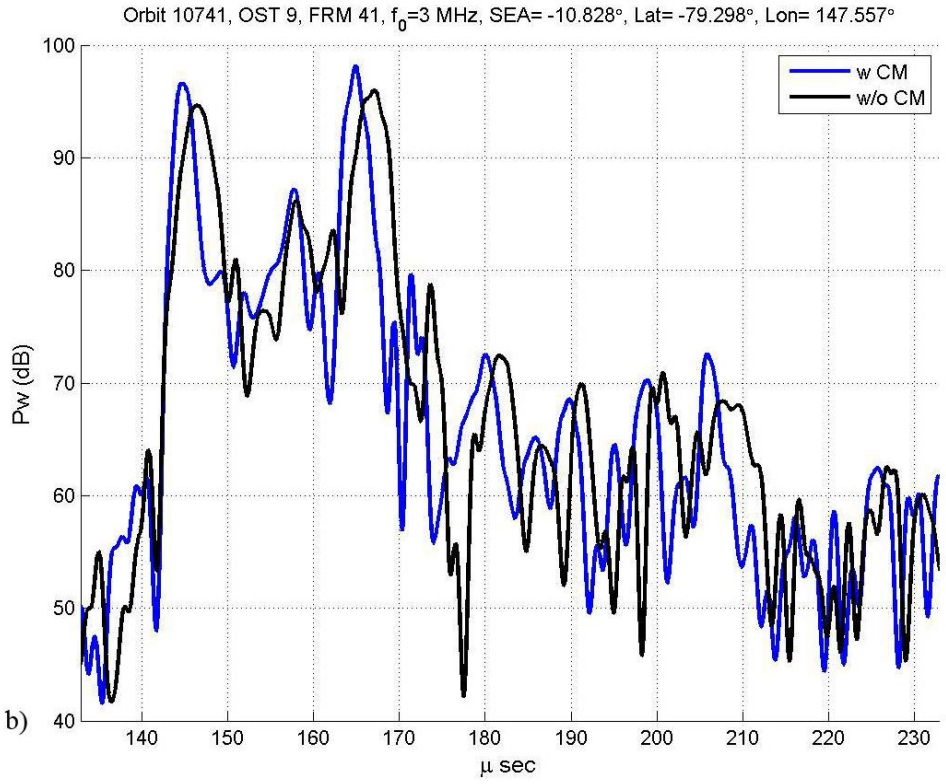
Fig. 6. SNR behavior, for 4 MHz, connected to the radargrams of Fig. 4. The blue and the red lines show the SNR obtained with and without the CM respectively.

From Fig. 5 we can see that the use of the CM (blue line) allows a gain in the SNR of 6 dB for 3 MHz and SEA around -7°, respect to the SNR obtained without the CM (red line). While from Fig. 6, the average gain, between the SNR obtained with the CM (blue line) and without the CM (red line), is around 5 dB with peaks of 8 dB for 4 MHz and SEA around 7°.

In details Fig. 7 shows the comparison between single frames selected from previous Fig. 3, that are collected with f_0 = 3 MHz and different SEA values and processed with (blue lines) and without (red lines) the CM. While in Fig. 7a, for SEA ≅ -13°, the differences are very light, in Fig. 7b, for SEA ≅ -11°, the main lobes of the signal without the CM become appreciable wider than the ones of the signal processed with the CM. Finally, in Fig. 7c, for SEA ≅ -7°, the chirp shape of the signal processed without the CM is completely compromised, with heavy losses in terms of range resolution and SNR. The different values along the abscissa of the three panels, are mainly due to the presence of the ionosphere delay, that, as anticipated at the end of section 4, it is not compensated by the CM; while the surface topography influence is limited to few μsec. Obviously the delay increases steadily passing from -13° to -7° of SEA.

The improvements, in terms of SNR, chirp shape and resolution, allowed by the CM increase with higher positive values of SEA, until the transmitted frequency it is too close to the plasma frequency. When the constraint $(f_{pmax} / f_0)^2 << 1$ is not verified, the CM is not able to compensate completely the ionosphere distortion, so the signal performances start to decrease.


Orbit 10741, OST 9, FRM 11, f0=3 MHz, SEA= -13.109°, Lat= -81.699°, Lon= 142.524°
w CM
w/o CM
Pw (dB)
μ sec
a)


Orbit 10741, OST 9, FRM 41, f0=3 MHz, SEA= -10.828°, Lat= -79.298°, Lon= 147.557°
w CM
w/o CM
Pw (dB)
μ sec
b)

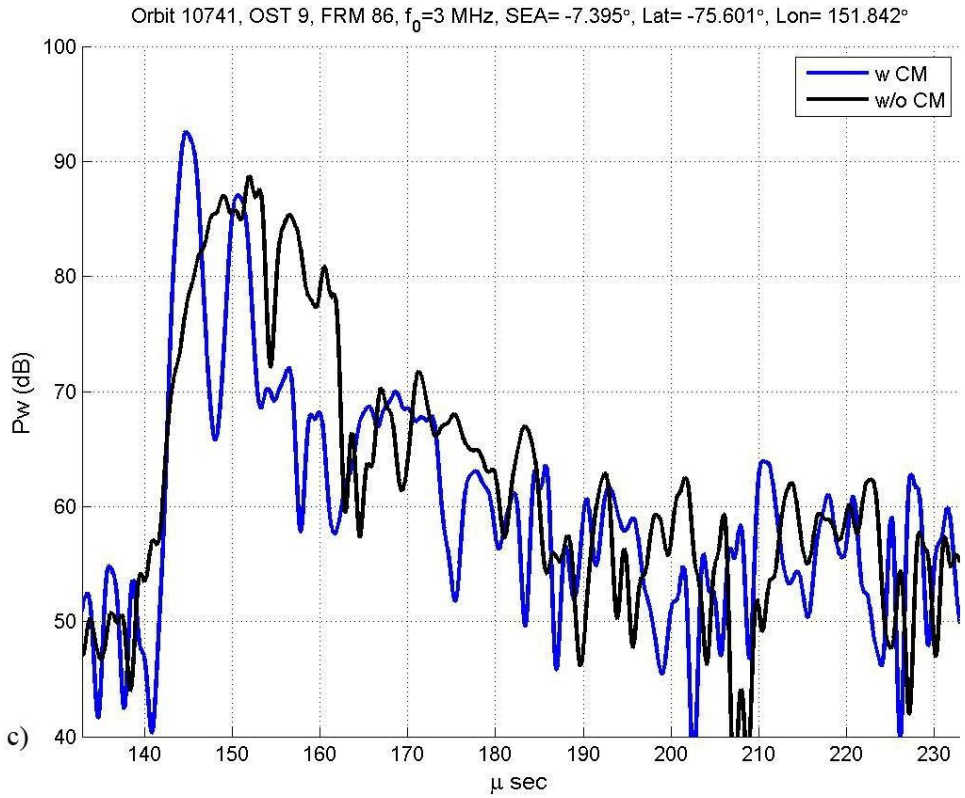


Fig. 7. The figures show the comparison between single frames collected with different SEA values (SEA ≅ -13° for Figs. 7a, SEA ≅ -11° for Figs. 7b, SEA ≅ -7° for Figs. 7c) and processed with (blue lines) and without (red lines) the CM.

6. Discussion and summary

In this document we have described in details the characteristics and the performances of the Contrast Method, a tool developed to compensate ionospheric distortion effects on radar signals. We have demonstrated that the data collected by the MARSIS radar in its subsurface mode, maintain their scientific information only if processed through the CM. The phase distortion introduced by the Mars ionosphere has proven to be an hard constraint for a radar sounder like MARSIS and its low operative frequencies, reducing considerably the key radar parameters (the SNR, the Side Lobes Level (SLL) and the range resolution) performances even during the sunset, when the SEA is still negative, but the upper portion of the ionosphere is already illuminated by the

sun. In these limitative environmental conditions, we have shown that the Contrast Method allows to optimize the radar performances not only during the night side, in particular the sunset, but also during the dayside, extending the MARSIS operations up to SEA ≅ 30° for f_0 = 5 MHz. It is evident, from Fig. 7, , that without the CM algorithm, the MARSIS operations would be limited only during the deep night side, i.e. SEA≤-15°÷20°, with a remarkable reduction of the coverage.

References

Cook C.E., Bernfeld M., Radar Signals, Academic Press, New York, 1967