DOT and SLA stationary and time-varying analytical covariance functions for LSC-based heterogeneous data combination

G.S. Vergos, D.A Natsiopoulos, I.N. Tziavos, V.N Grigoriadis and E.A Tzanou

Department of Geodesy and Surveying, Aristotle University of Thessaloniki, Univ. Box 440, 54 124, Greece

Introduction and Problems

With the availability of an abundance of earth observation data from satellite altimetry missions as well as those from the GOCE satellite, monitoring of the sea level variations and the determination of functionals of the Earth's gravity field are gaining increased importance.

One of the main issues of heterogeneous data combination with stochastic methods is the availability of appropriate data and error covariance and cross-covariance matrices.

The latter needs to be determined for all input data within an LSC-based combination scheme based on some analytical global covariance function models, which interconnect observations and signals to be predicted.

Given the availability of altimetric sea surface heights, GOCE observations of the second-order derivatives of the Earth's potential, geoid height variations from GRACE and marine gravity anomalies, one can employ all such available information within LSC to estimate the mean dynamic ocean topography (DOT) as well as its dynamic, i.e., time-varying part.

Data used and corrections

The focus is based on three different models of DOT (2D case). The first is the model of Rio and Hernadez (2004) the second is the model of Rio et al. (2007) and the third one is a combined (GOCE, GRACE and altimetry) model developed in the frame of GOCESeaComb.

The Rio and Hernadez model is an MDT computed over the entire world from altimetry, in situ measurements, and a geoid model.

From available data only the values of the MDT in the Mediterranean Sea have been used. These are 15557 values in $1/8^{o}$ grid spacing and are bounded between $30.75° \leq \phi \leq 45.625°$ and $-5° \leq \lambda \leq 35.75°$.

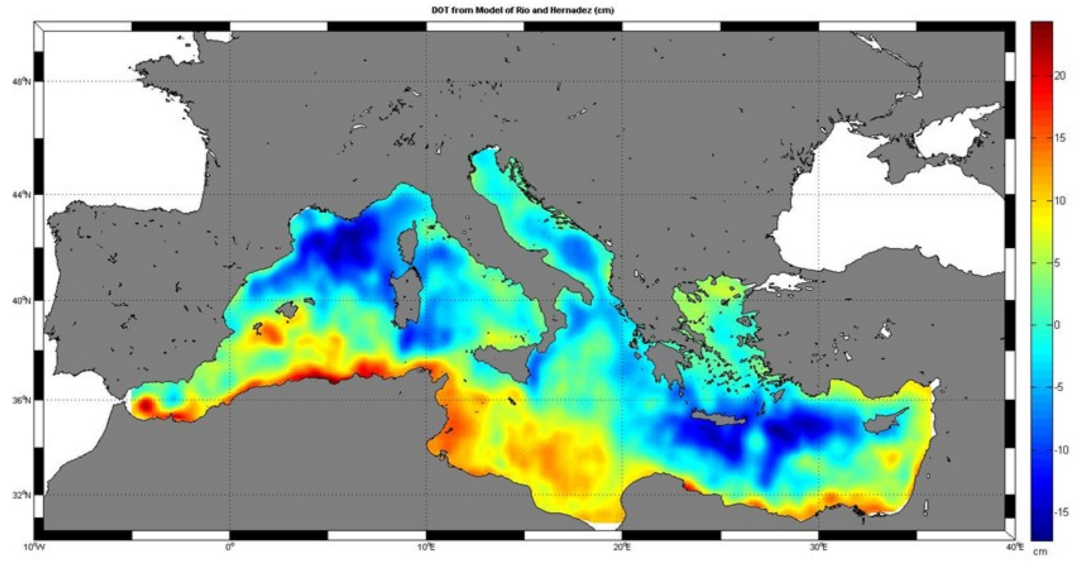
The Rio et al. model is an MDT model of the Mediterranean Sea computed from altimetric data, in-situ measurements and a general circulation model. These are 32435 values, spanning the entire Mediterranean Sea and bounded between $30° \leq \phi \leq 50°$ and $-10° \leq \lambda \leq 40°$.

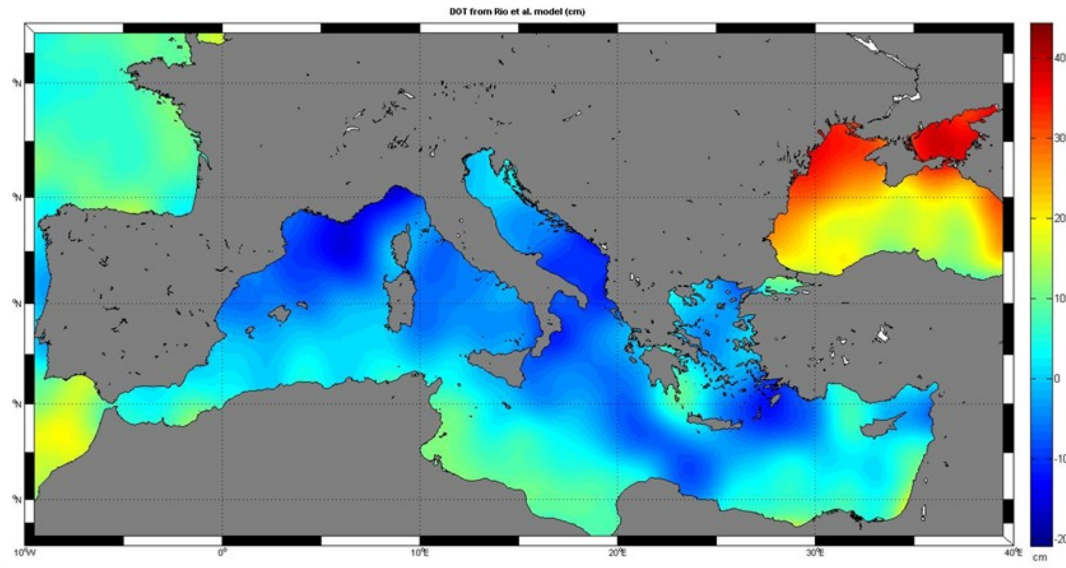
The combined model was derived from a combination of a GGM and the DTU2010 MSS model. The MDT was derived as a deviation from the MSS model, which was then filtered in order to treat high-frequencies, as well as the omission and commission errors. In this part of the work the one to be used refers to boxcar filtering function for a cut-off wavelength of λ=200 km.

$$h(x,y)=2\lambda_c\, sinc\left(2\lambda_c\left(x^2+y^2\right)\right) \qquad H(u,v)=\prod\left(\frac{\omega}{2\omega_c}\right)$$

The combined model consists of 55639 values in 5' x 5' grid spacing that span the entire Mediterranean Sea bounded between $30° \leq \phi \leq 50°$ and $-10° \leq \lambda \leq 40°$.

For the time-varying MDT analytical covariance functions, a 1D case was studied using pass 196 from JASON-1 and pass 444 form ENVISAT.





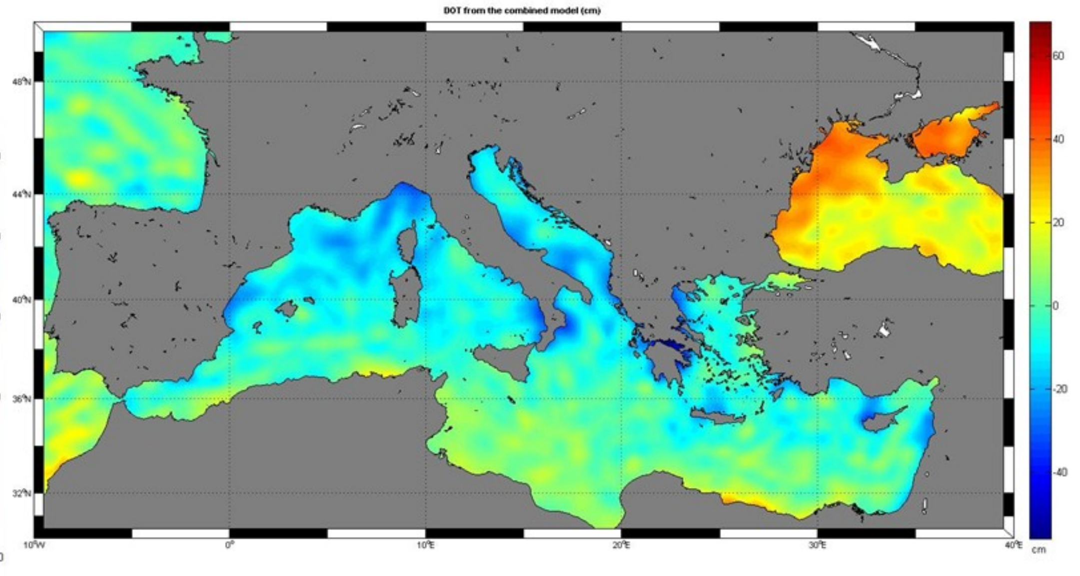


Figure 1:MDTs from the Rio and Hernadez model (left top), Rio et al. model (left bottom) and the combined GOCE/GRACE model (right).

Objectives

This work presents some news ideas and results on the determination of analytical covariance functions and subsequently full variance-covariance matrices for the DOT in the Mediterranean Sea.

Along track records of the DOT have been used to represent as close as possible the time-variable nature of the time-variable nature of the non-stationary DOT signal after estimating time varying analytical covariance functions

The estimation of the analytical covariance functions is performed using 2nd and 3rd order Markov models. The same analysis have been carried out for the time-varying functions.

A kernel model similar to that of the disturbing potential a.k.a dependent on a series of Legendre polynomials has been tested for an entire window both for the three models of DOT and for the whole model of Rio and Hernadez.

The goal is to come to some conclusions on the stationary and time-varying DOT spectral characteristics based on the empirically derived properties such as the variance and correlation length and determine analytical models to be used later for prediction with LSC.

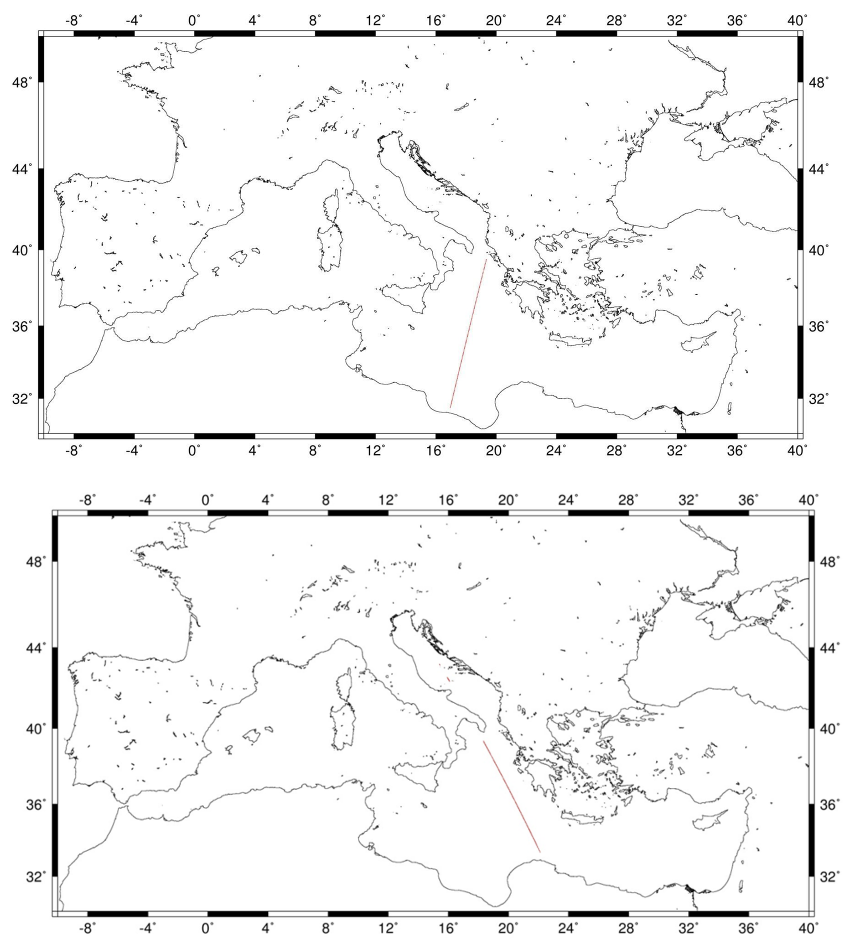


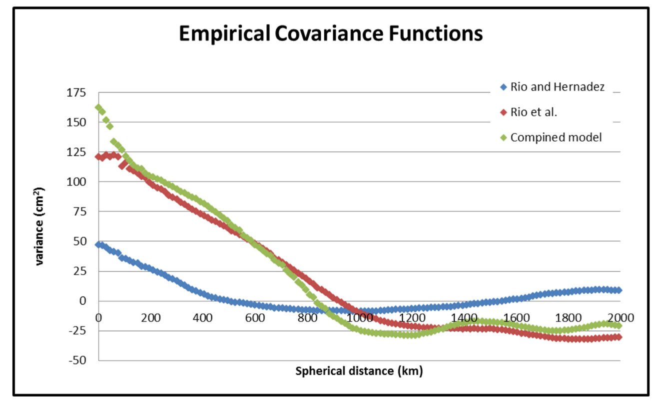
Figure 2: ENVISAT pass 444 (top) and Jason pass 196 (bottom).

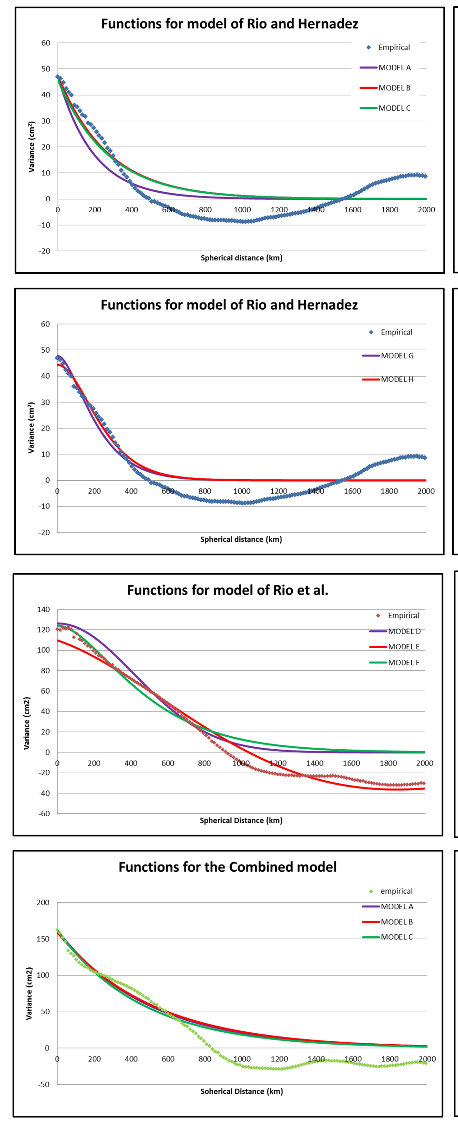
Mathematical models and Covariance estimation

First the empirical covariance functions have been estimated for the three DOT models, the variance C_0 the correlation length ξ were determined.

Then, various analytical covariance function models have been investigated in order to determine the one that provides the overall best fit to the empirical model as well as the optimal results, in terms of prediction accuracy. To this extend, various order exponential models have been studied, along with second and third order Gauss- Markov ones.

Apart from planar models, a spherical one based on Legendre polynomial expansion, simulating the Tscherning & Rapp model used to model the analytical covariance function of the disturbing potential was used.





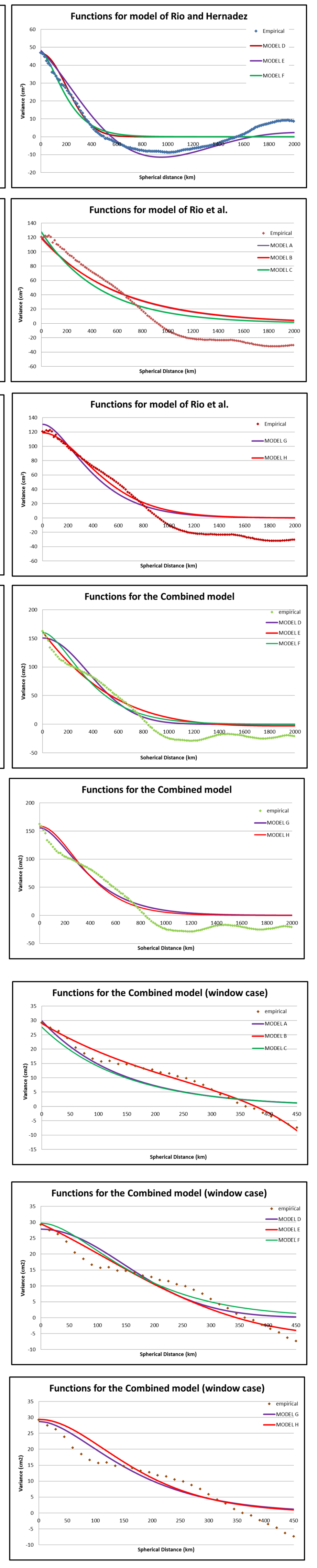


Figure 3: Empirical and analytical covariance functions for all types of DOT models

Exponential Models

$$C(\psi)=\alpha e^{b\psi} \quad \textbf{Model A}$$

$$C(\psi)=\alpha e^{b\psi}+ce^{d\psi} \quad \textbf{Model B}$$

$$C(\psi)=\alpha e^{-\left(\frac{\psi-b}{c}\right)^2} \quad \textbf{Model C}$$

$$C(\psi)=\alpha e^{-b\psi^2} \quad \textbf{Model D}$$

$$C(\psi)=\alpha e^{-b\psi}\cos(\omega\psi) \quad \textbf{Model E}$$

$$C(\psi)=\alpha(1+b\psi)e^{-b\psi} \quad \textbf{Model F}$$

Gauss-Markov Models

$$C(r)=\sigma(\bullet)^2\left(1+\frac{r}{D}\right)e^{\left(\frac{-r}{D}\right)} \quad \textbf{Model G}$$

$$C(r)=\sigma(\bullet)^2\left(1+\frac{r}{D}+\frac{r^2}{3D^2}\right)e^{\left(\frac{-r}{D}\right)} \quad \textbf{Model H}$$

Spherical model based on Legendre polynomial expansion

$$C_{(\bullet,\bullet)}(\psi)=\sum_{n=0}^{\infty}\sigma(\bullet)_n^2 P_n(\cos\psi) \qquad \sigma(\bullet)_n^2=b\left(\frac{k_2^3}{k_2^3+n_3}+\frac{k_1^3}{k_1^3+n^3}\right)s^{n+1} \quad \textbf{Model I}$$

In the above models, ***ψ*** denotes the spherical distance, ***ξ*** the correlation length, ***r*** the planar distance and ***σ(•)²*** the variance of quantity ***(•)*** which is the investigation (SLA or DOT). The rest, are parameters to be determined, so that the analytical model will fit the empirical one. Note, that for all models a mixed equations adjustment scheme was used in order to determine the necessary parameters for each model, based on the empirical values.

Prediction has been carried out by omitting every second point where values of ζ are available using the rest for the estimation. This was performed for the two DOT models of Rio, for all available data, while two window cases have been tested also.

Due to the big dimensions of the matrices, prediction for the model of Rio et al. has been made using the method of singular value decomposition for the inversion of big matrices. This method has also been applied to model D for Rio and Hernadez and to the combined model (window case) due to problems in the inversion of the matrix (ill-posed matrices). In the window cases for the model of Rio and Hernadez the total points were 2397 and the values predicted were 1199 while for the combined model the total points were 4857 and the values predicted were 2479.

Statistics of Rio and Hernadez (cm)				
	min	max	mean	std
ζ	-17.183	24.357	0.854	±6.851
Prediction errors with LSC for the various covariance models (cm)				
MODEL A	-0.944	2.220	0.002	±0.159
MODEL B	-1.021	2.099	0.001	±0.156
MODEL C	-1.372	1.442	-0.001	±0.149
MODEL D	-6.867	4.630	-0.004	±0.938
MODEL E	-1.140	1.797	0.000	±0.152
MODEL F	-0.988	0.842	-0.002	±0.087
MODEL G	-0.988	0.842	-0.002	±0.087
MODEL H	-1.542	0.852	-0.002	±0.085
Rio et al. (cm)				
ζ	-20.600	44.300	4.379	±10.988
Prediction errors with LSC for the various covariance models (cm)				
MODEL A	-1.699	1.101	0.000	±0.052
MODEL B	-1.700	1.010	0.000	±0.052
MODEL C	-0.984	1.436	0.000	±0.052
MODEL D	-8.141	4.890	0.003	±0.892
MODEL E	-0.990	0.700	0.000	±0.043
MODEL F	-0.300	0.246	0.000	±0.038
MODEL G	-0.305	0.246	0.000	±0.038
MODEL H	-0.300	0.245	0.000	±0.038
Combined model (window case) (cm)				
ζ	-19.650	13.350	1.723	±5.123
Prediction errors with LSC for the various covariance models (cm)				
MODEL A	-1.482	0.776	-0.001	±0.090
MODEL B	-1.519	1.050	-0.006	±0.106
MODEL C	-1.421	0.641	-0.001	±0.086
MODEL D	-9.840	30.067	-0.065	±1.850
MODEL E	-1.433	0.856	-0.001	±0.090
MODEL F	-0.296	0.411	0.000	±0.036
MODEL G	-0.296	0.410	0.000	±0.036
MODEL H	-0.350	0.434	0.000	±0.031
Rio and Hernadez (window case) (cm)				
ζ	-9.200	10.600	1.125	±5.557
MODEL I (a)	-13.169	0.932	-0.007	±1.556
MODEL I (b)	-0.526	0.932	0.193	±0.055

Table 1: Statistics from stationary DOT models and prediction errors from all models.

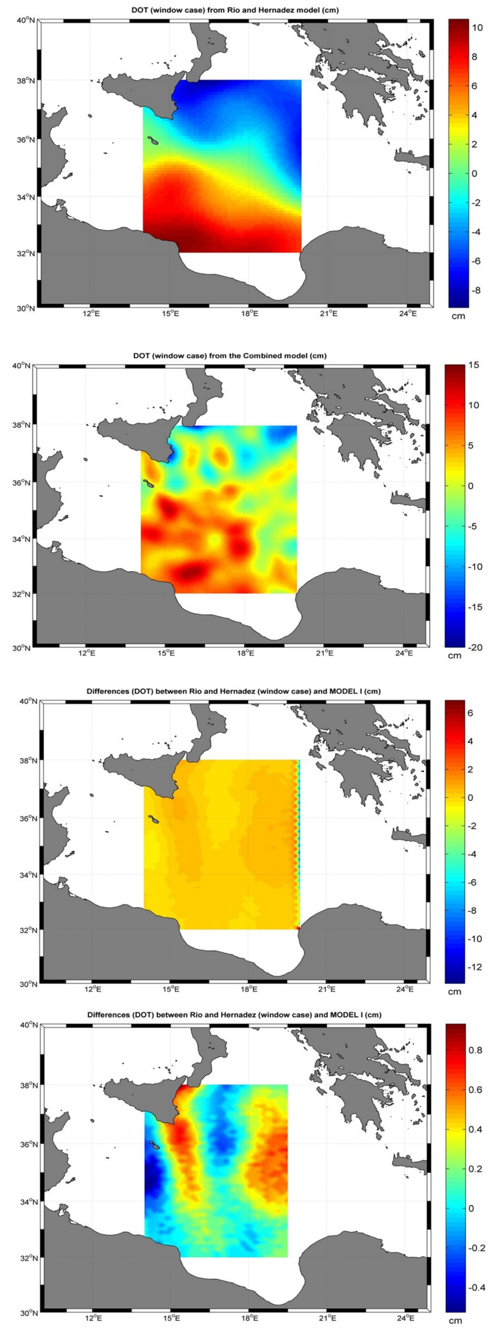


Figure 4: DOT from Rio et Hernadez (1), DOT from the combined model (2), differences between Rio and Hernadez and MODEL I(a) (3), differences between Rio and Hernadez and MODEL I(b) (4).

Time-varying DOT and pre-processing

Data about the time-varying DOT (Δζ) topography that have been derived from the combination of altimetric records (SLAs) of Envisat and Jason1 and the Rio et al. model of DOT .

$$\Delta\zeta = SSH-N-\zeta \qquad \Delta\zeta = SLA-\zeta$$

Both passes, Envisat 444 and Jason 196, have been selected based on the following criteria: a) both passes shall be in almost the same area b) the passes shall be long and span the entire basin in the north-south or south-north direction (ascending or descending pass respectively), c) there shall be no or little land intrusion from isles or islands in the pass SLA records, d) the data record shall be as consistent as possible throughout the satellite data record for the period of study, i.e., missing records and/or voids should be kept to a minimum.

From all available data those of year 2009 have been selected. Jason-1 period is 10 days (actually 9.9 days) and Envisat one is 35 days. As a result Jason-1 data consist of 34 cycles and a total amount of 5849 point values. On the other hand, Envisat data consist of 10 cycles and 1206 point values.

To estimate *Δζ, ζ* values should be calculated in points where SLA data are available. MDT was interpolated at the along-track locations by LSC using the covariance functions already determined. Moreover, the next step of pre-processing was to have the same number of points per cycle for Envisat and the same number of points per cycle Jason-1.

Cycles and total points per cycle for ENVISAT

cycle	total points
76	132
77	123
78	123
79	118
80	102
81	123
82	121
83	123
84	119
85	122

Cycles and total points per cycle for JASON-1

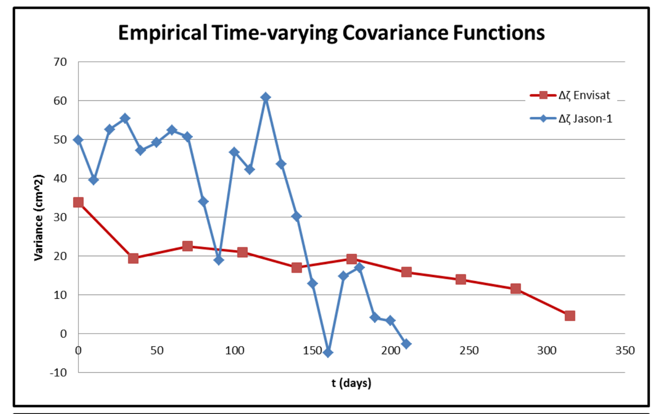
cycle	total points	cycle	total points	cycle	total points
19	139	34	147	48	228
20	148	35	146	49	166
21	153	36	151	50	194
22	151	37	163	51	230
23	147	38	231	52	152
24	153	39	204	53	147
27	123	40	141	54	138
28	131	41	201	55	226
29	201	42	230		
30	153	44	227		
31	117	45	226		
32	146	46	163		
33	149	47	227		

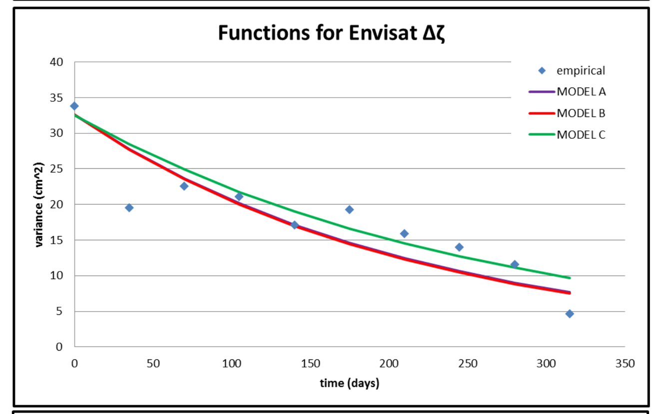
The big differences in the number of points in Jason-1 cycles are due to the fact that some cycles have SLA values over northern Italy.

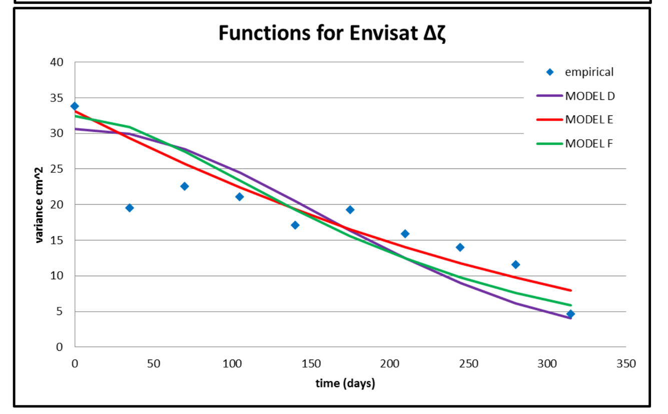
From all available points only those over the same location, within Δφ=0.01° and Δλ=0.01°, have been selected, since they refer to the same locations.

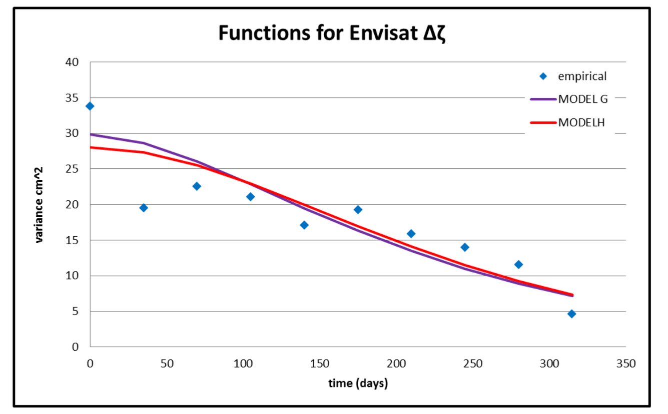
As a result the number of collocated points for Envisat passes was 93 (total points=93*10=930) and 129 points for Jason-1 (all points north of Italy have been excluded-total points=129*37=4773).

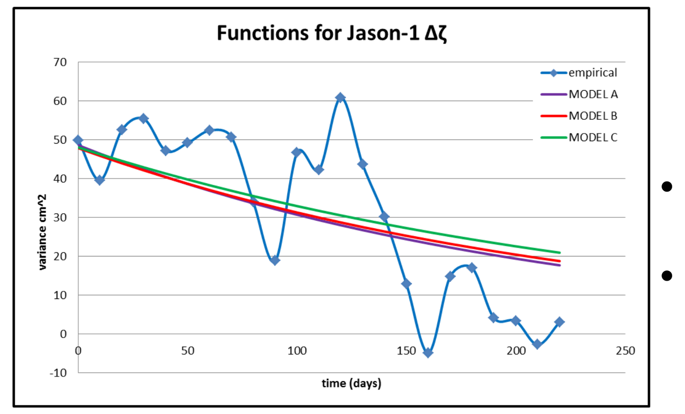
Models A to H have been used to determine analytical covariance functions for the time-varying DOT. In all models spherical distance **ψ** has been replaced by time **t.**











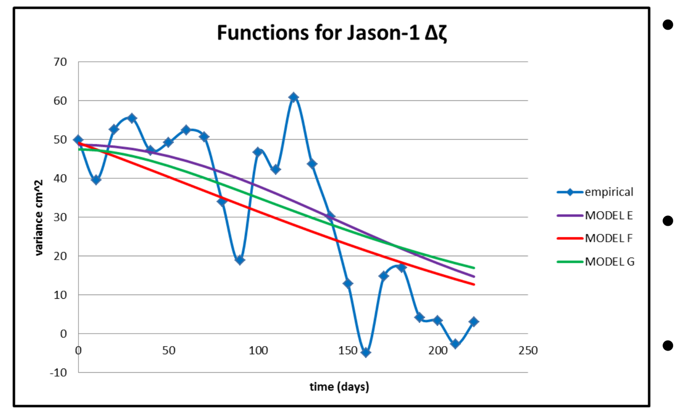


Figure 5: Empirical and analytical covariance functions for time-varying DOT.

For the Envisat-derived *Δζ*, all models provide good fit to the empirical values. This is logical as the form of empirical covariance function is close to this of an exponential model and as a result all models fit well to the empirical. On the other hand, for Jason-1, the miss-fit of the models in the biggest part of the equation is obvious due to the badly scaled covariance function.

Prediction has been carried out by omitting every second point where values of *Δζ* are available using the rest for the prediction.

Prediction has been made for two models of *Δζ* for all available data. Due to the badly scaled and close to singular matrices, prediction for the model of Jason-1 with all models has been made using the method of singular value decomposition for the inversion of matrices.

This method has also been applied to model D and model F for Envisat model due to problems in the inversion of the matrix (ill-posed matrices).

Statistics of Envisat Δζ				
	min	max	mean	std
Δζ	-27.1	26.4	3.4	±10.3
MODEL A	-8.200	10.497	0.165	±2.299
MODEL B	-8.200	10.497	0.165	±2.299
MODEL C	-8.200	10.497	0.165	±2.299
MODEL D	-19.971	19.694	0.129	±6.329
MODEL E	-8.200	10.497	0.165	±2.299
MODEL F	-22.018	14.438	0.139	±5.521
MODEL G	-8.200	10.497	0.165	±2.299
MODEL H	-8.200	10.497	0.165	±2.299

Statistics of Jason-1 Δζ				
	min	max	mean	std
Δζ	-36.024	32.367	5.325	±11.161
MODEL A	-19.032	11.051	0.010	±2.033
MODEL B	-19.032	11.051	0.010	±2.033
MODEL C	-19.032	11.051	0.010	±2.033
MODEL D	-30.693	27.789	0.027	±9.025
MODEL E	-19.032	11.051	0.010	±2.033
MODEL F	-16.479	25.407	0.015	±4.606
MODEL G	-19.032	11.051	0.010	±2.033
MODEL H	-16.479	25.407	0.015	±4.606

Table 2: Statistics of the time-varying DOT models and prediction errors with the various analytical models. [Units: mm]

Conclusions

- For the time-varying DOT, all models manage to give small errors, of the order of 2 mm.
- This can be attributed to the existence of many data in the region where predictions need to be made and to the small differences on the time of nearby observations.
- On the other hand, the small errors with std of the order of a few mm indicate that the determined analytical covariance functions perform well, so that the estimates determined are rigorous and robust.
- For the stationary DOT, all analytical models, except model D, present very small errors as well, of the order of 1-2 mm.
- The small errors with std of the order of a few mm indicate that the determined analytical covariance functions perform well, so that the estimates determined are rigorous and robust.
- The exponential model F and the two Gauss-Markov models give the best results with the standard deviation of the prediction errors at the ±0.085 cm.
- For MODEL I when the whole region is concerned (Ia), there are some edge effects in the eastern part of the area.
- This is shown in the predicted field with minimum values of the order of -13 cm compared to –9 cm for the original DOT.
- On the other hand when the entire window used is 0.5° smaller (b) (19.5°-20°) blunders are removed and the standard deviation of the field is 1.5 cm smaller.
- The next step refers to the determination of 2D time-varying DOT models for the entire Mediterranean through dedicated spatio-temporal covariance function for the entire duration of the ENVISAT and JASON-1 missions.

EGU General Assembly, Session G1.1 – "Recent Developments in Geodetic Theory", EGU2014-1040, B764
27 April - May 02 2014, Vienna – Austria