Satellite Altimetry Based Absolute Dynamic Topography Prediction Using Machine Learning in Baltic Sea

Majid Mostafavi[1] , Nicole Delpeche-Ellmann[2] , Artu Ellmann[1]

[1] Department of Civil Engineering and Architecture, Tallinn University of Technology (majid.mostafavi@taltech.ee), [2] Department of Cybernetics, Tallinn University of Technology

Introduction

Determining Absolute Dynamic Topography (DT) of the ocean is a key component in examining realistic sea level and for understanding meso-scale dynamics. It is common that tide gauge (TG) data once referred to the geoid can provide high resolution DT estimates however, limited to a fixed location in the coastal areas. Satellite altimetry (SA) provides sea level data along their tracks both at the coast and offshore. Whilst SA has some limitation (e.g., spatial, and temporal limitations, in adequate corrections, land contamination etc.), utilizing high-resolution geoid models allows the determination of a more realistic DT. This study examines different techniques with SA data to predict DT at a temporal resolution of 24 months.

Multiple autoregression methods including Autoregression **(AR)**, Moving Average **(MA)**, Autoregressive Moving Average **(ARMA)**, Autoregressive Integrated Moving Average **(ARIMA)** and Seasonal Autoregressive Integrated Moving-Average **(SARIMA)** are employed using SA data during 1995-2019 (that includes ERS-2, Envisat, SARAL, Jason-1, Jason-2, CryoSat-2, Jason-3, Sentinel-3A and Sentinel-3B) to predict detrended absolute DT. The predicted SA data are evaluated against the 12 TG stations records in Baltic Sea. The results of this study of a more realistic DT can be applicable for engineering, navigation, and coastal management. Also, careful spatial data selection and outliers' removal using data screening is prerequisite.

Datasets

TG: 12 TG stations data in Baltic Sea (Table 1)

SA: Sentinel-3A (S3A), Sentinel-3B (S3B), SARAL (SRA), Envisat (ENV), ERS-2 (ER2) data during 1995-2019 using ALES+ retracker (Baltic+SEAL)

Model: NKG2015 Geoid & NKG2016LU Land Uplift

Objectives

- ❑ Detecting most robust model for SA data times series forecasting at vicinity of TG station.
- ❑ Sea Level Trend analysis using forecasted SA data in the Baltic Sea.

ID	TG Station		Latitude [°]	Longitude [°]
1	Helsinki	(FI)	60.15363	24.95622
2	Hamina	(FI)	60.56277	27.1792
3	Hanko	(FI)	59.82287	22.97658
4	Pietarsaari	(FI)	63.70857	22.68958
5	Kaskinen	(FI)	62.34395	21.21483
6	Rauma	(FI)	61.13353	21.42582
7	Oulu	(FI)	65.0403	25.4182
8	Spikarna	(SW)	62.3633	17.5311
9	Marviken	(SW)	58.5536	16.8372
10	Simrishamn	(SW)	55.5575	14.3578
11	Forsmark	(SW)	60.4086	18.2108
12	Furuögrund	(SW)	64.9158	21.2306

Table 1. TG stations list

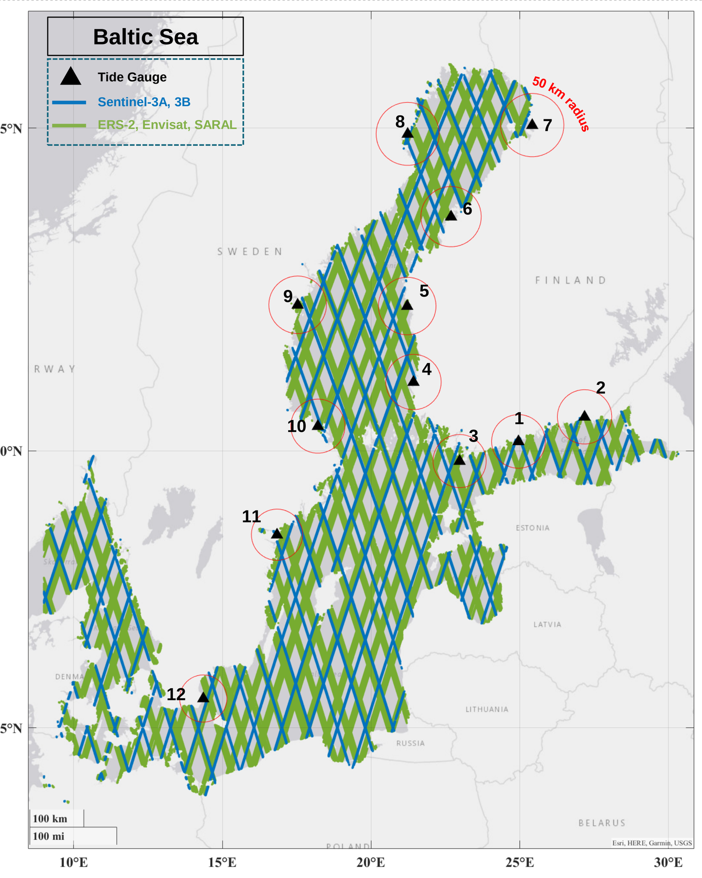


Fig 1. Location of TG stations and SA passes in BS

Methodology

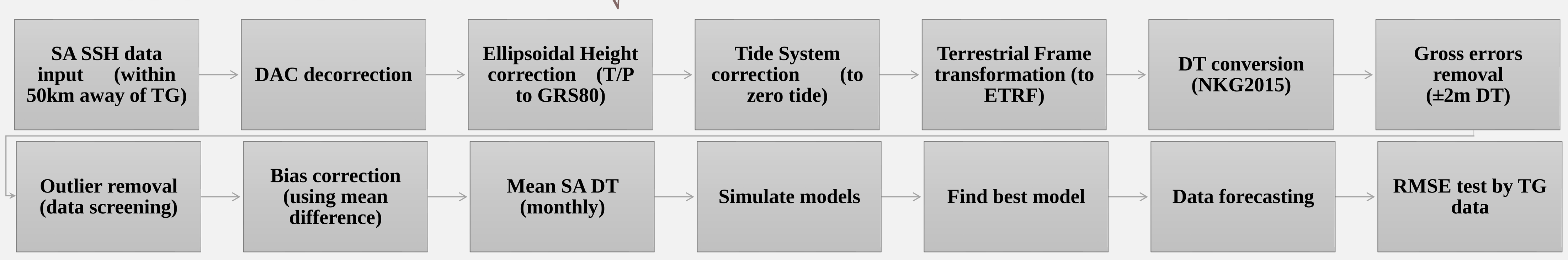
1. Spatial selection and outlier detection: extract SA data within 50km from TG and screen out low-quality data using data screening.
2. Vertical land movement (VLM) correction for TG time series.
3. Bias between SA and TG applied as the difference of mean of each dataset at single TG location.
4. Monthly average of SA DT of each mission at near each TG considered as DT_{SA}.
5. Test DT_{SA} data stationary using Augmented Dickey-Fuller test. Also, transforming data to 1) detrended: reduce 1st degree polynomial trend of DT_{SA}, 2) difference: differences between adjacent elements of detrended DT_{SA}
6. SA DT time series forecasting for 2 years using different models including: Autoregression (AR), Moving Average (MA), Autoregressive Moving Average (ARMA), Autoregressive Integrated Moving Average (ARIMA), Seasonal Autoregressive Integrated Moving-Average (SARIMA).
7. Assess model adequacy using Akaike information criteria (AIC) and Bayesian (Schwarz) information criteria (BIC) to select the most robust model also using statistical evaluation criteria, including the correlation coefficient (R), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE):

$$R = \left[\frac{\sum_{i=1}^{n}(y_i^o - \overline{y^o})(y_i^r - \overline{y^r})}{\sum_{i=1}^{n}(y_i^o - \overline{y^o})\sum_{i=1}^{n}(y_i^r - \overline{y^r})}\right] \quad RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i^o - y_i^r)} \quad MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i^o - y_i^r|$$

n: number of data

y^o: observed DT_{TG}

y^r: forecast DT_{SA}



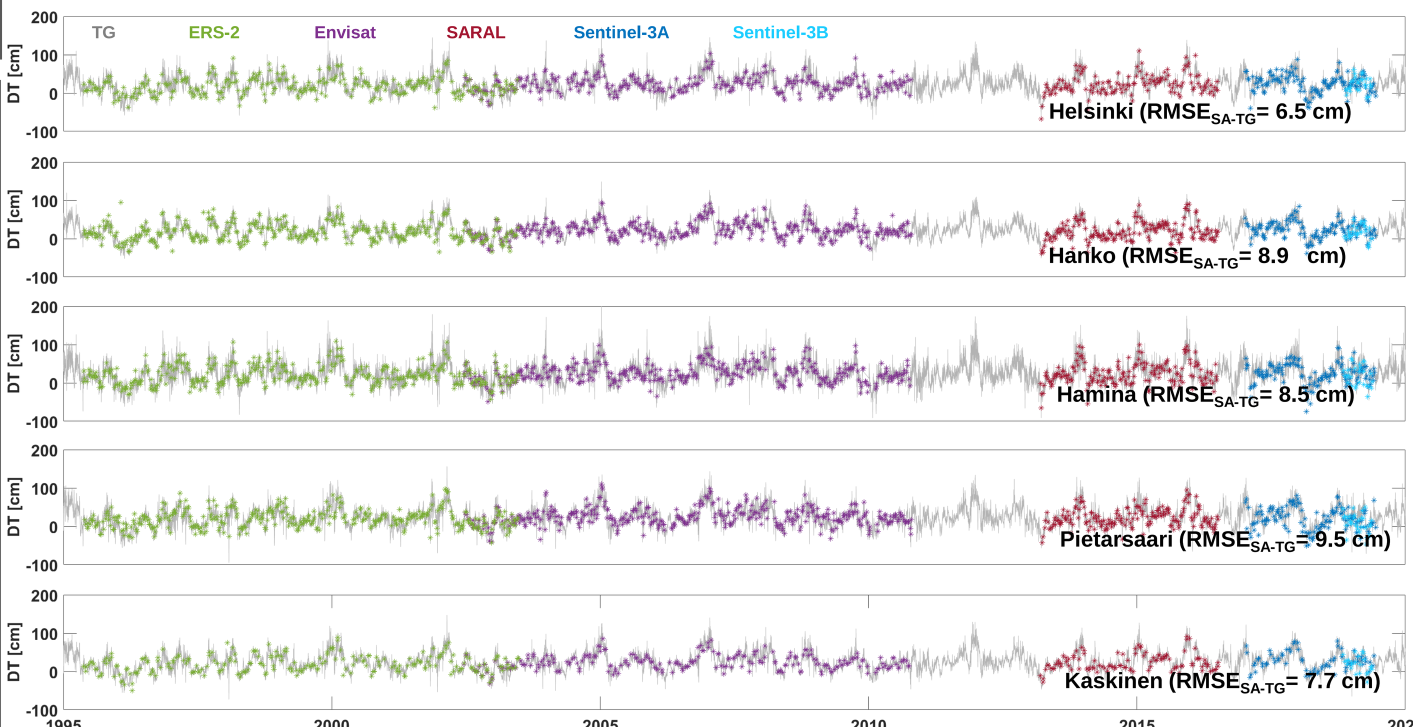


Fig 2. Sea Level changes at each TG station using TG and SA data

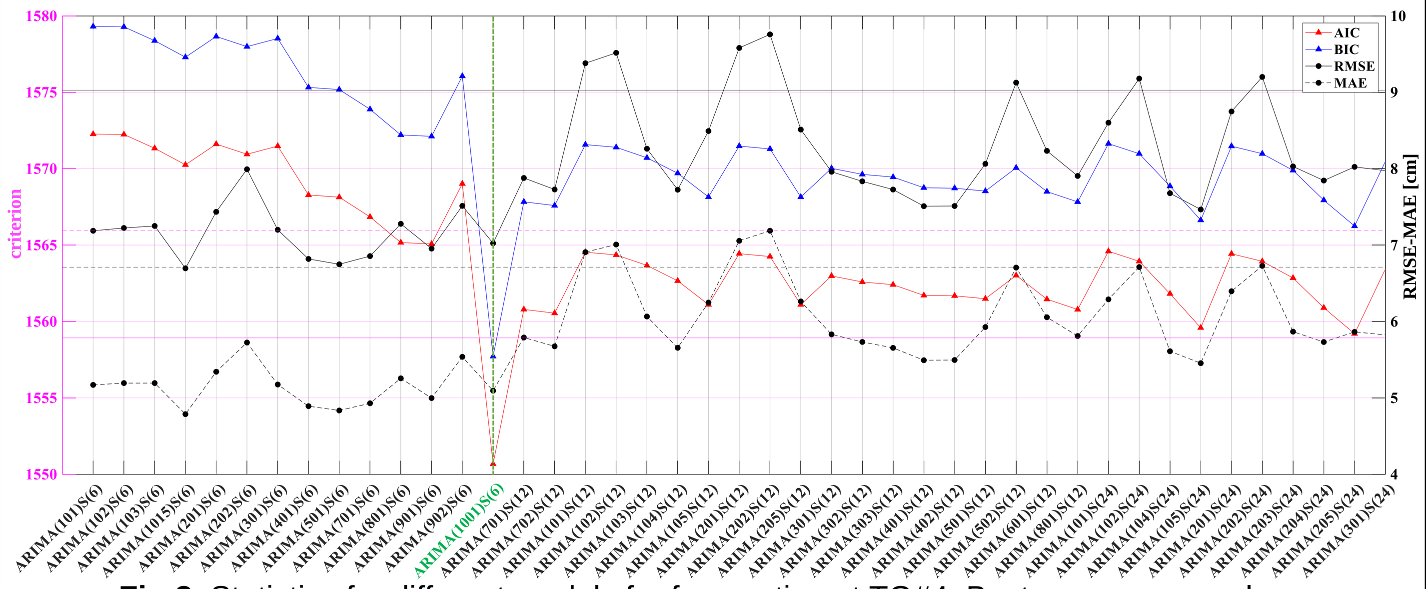


Fig 3. Statistics for different models for forecasting at TG#4. Best case: green colour

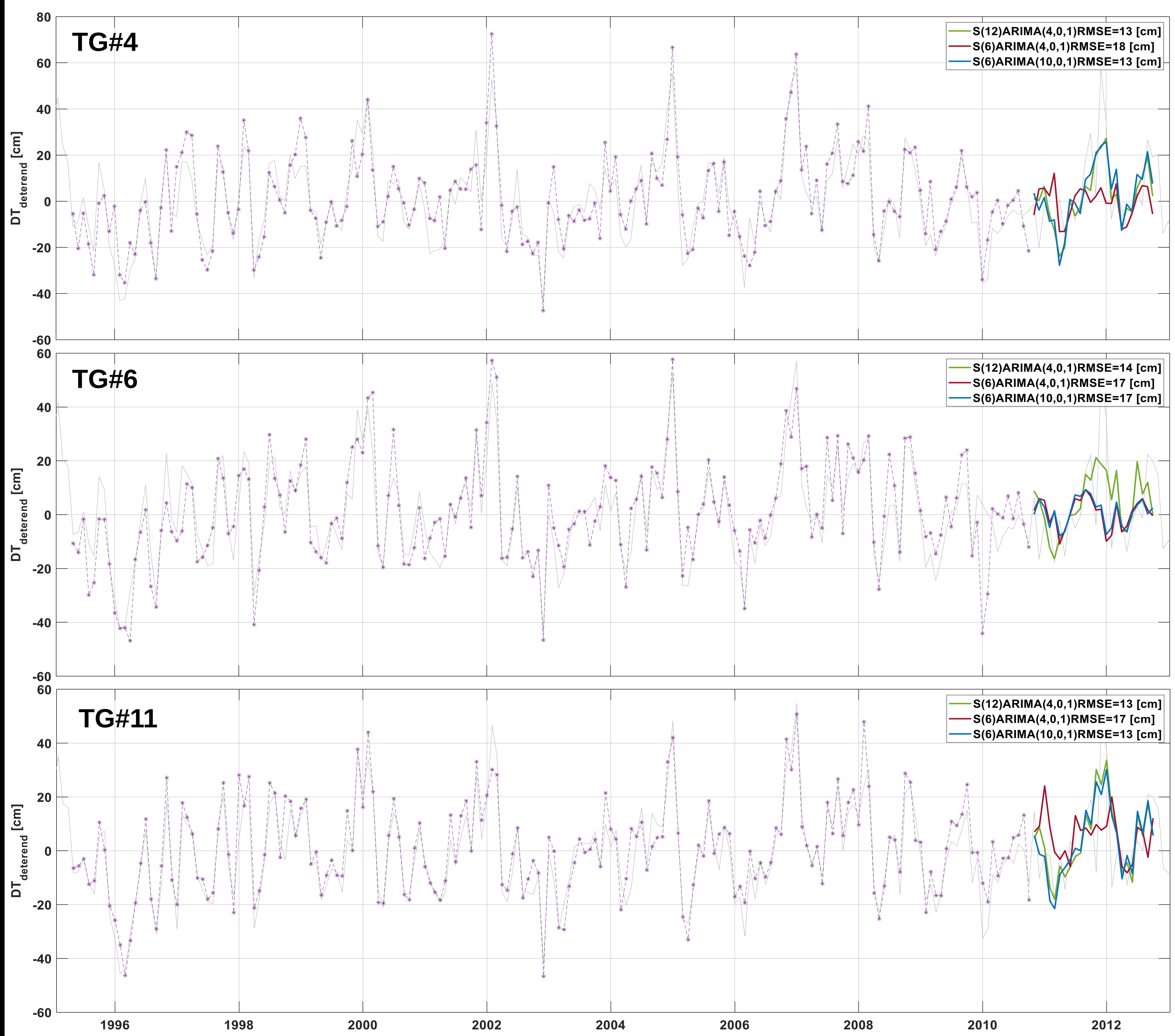


Fig 6. Sea Level forecasting using SA data (purple starts) and tree models at tree TGs (ID: 4,6 and 11), comparing to TG data (RMSE)

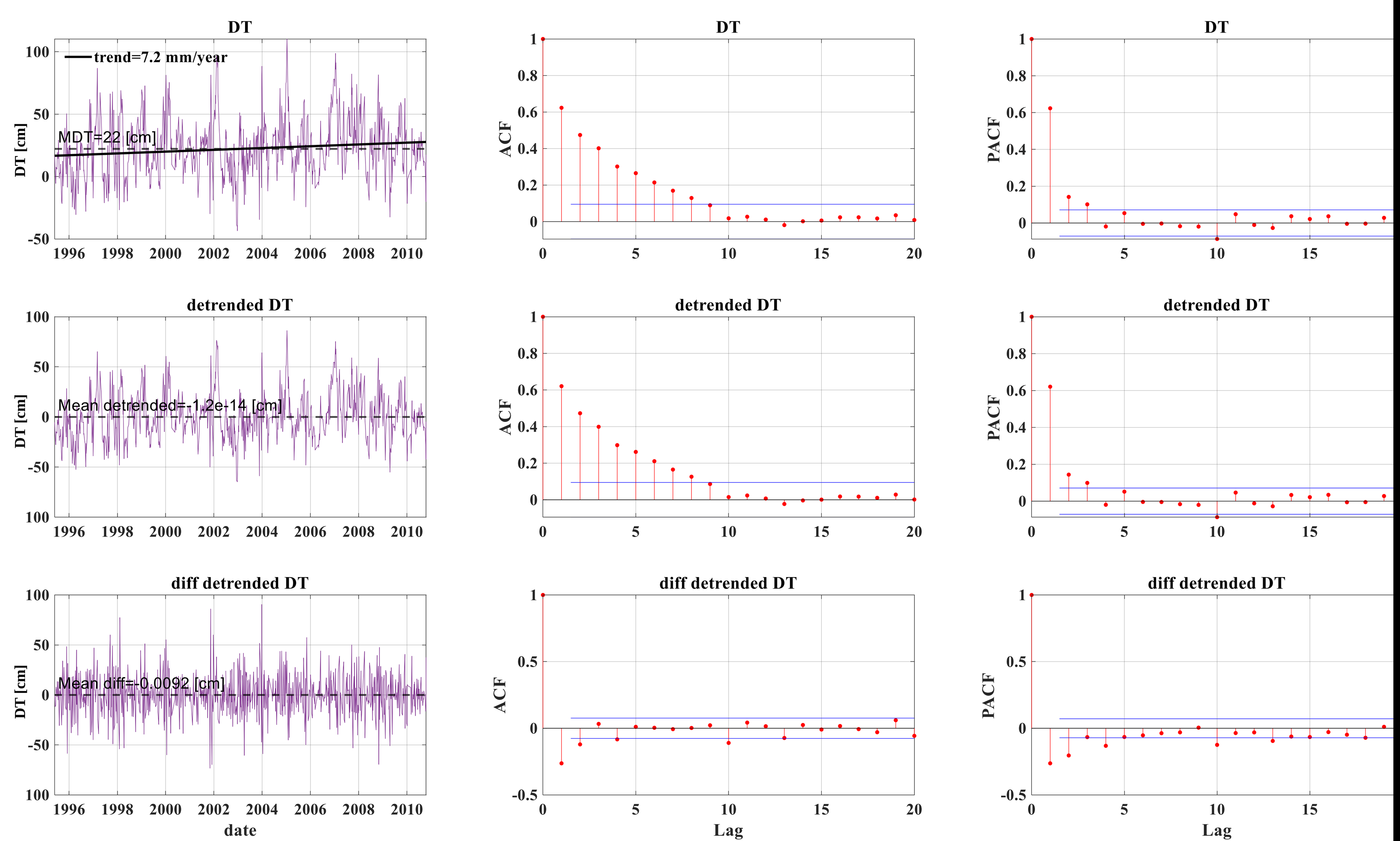


Fig 4. DT_{SA} ACF and PACF in tree cases: DT, DT detrended, difference of DT detrended at TG#4.

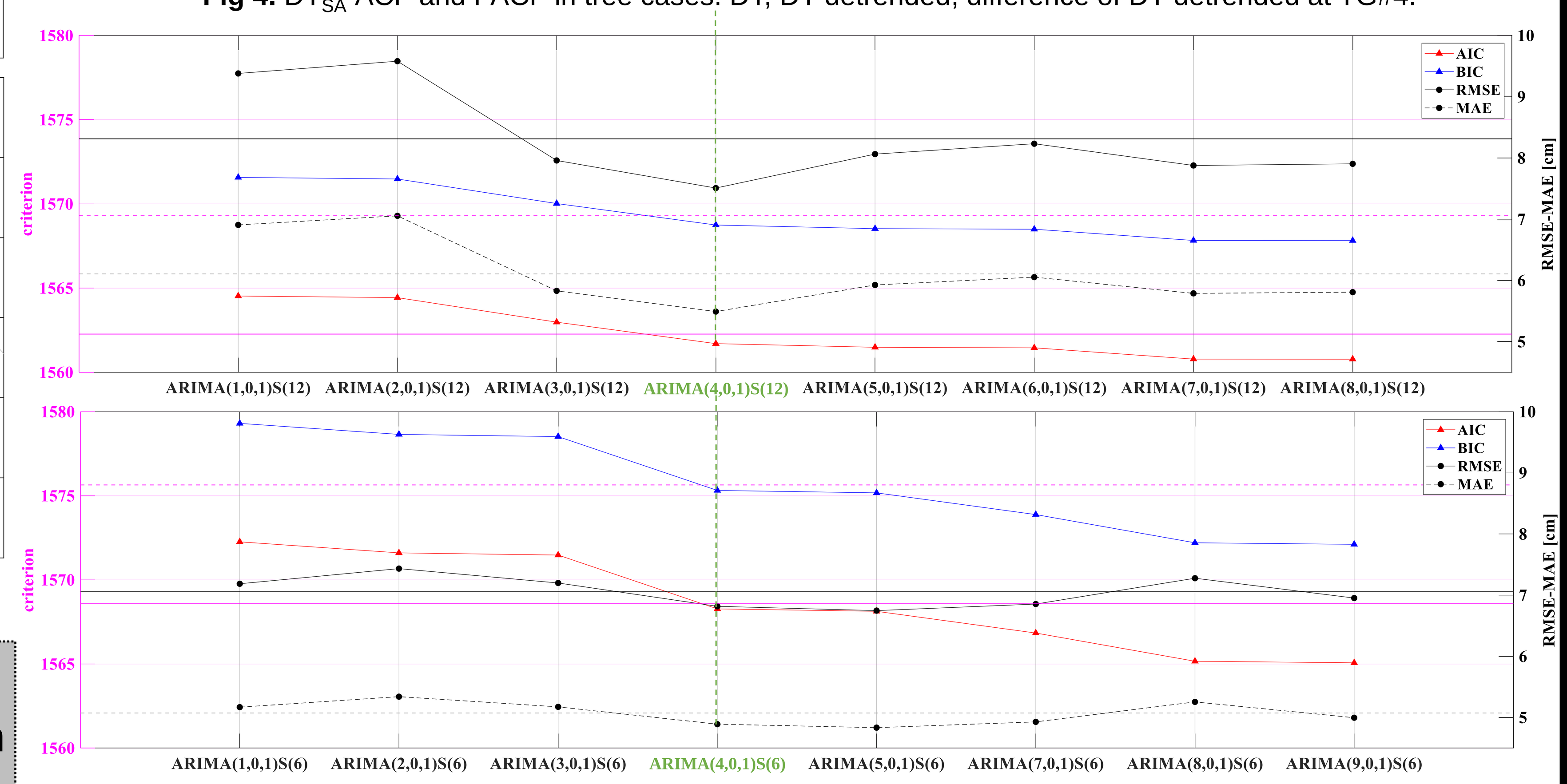


Fig 5. SARIMA models of DT_{SA} at TG#4 by two different seasonality (6 and 12 month). Best case: green

Results

- ► Good agreement is obtainable using SARIMA by ARIMA(4,0,1) SAR(48) SMA (12) with 12 month Seasonality model for DT_{SA} forecasting (Figure 3 and 5).
- ► DT data need to be detrended to consider for Sea Level Trend before forecasting (Figure 4).
- ► DT forecasting can be forecasted by ~15 cm RMSE using SA data for 2 years. This forecasted data can be served for 2 years data gap of SA data (during 2011-2013).

Acknowledgements

The research is supported by the Estonian Research Council grants PRG330 'Development of an iterative approach for near-coast marine geoid modelling by using re-tracked satellite altimetry, in-situ and modelled data'