Comparison of techniques for raster time series analysis

Filipponi Federico

ISPRA - Italian Institute for Environmental Protection and Research, Via Vitaliano Brancati 48, Roma (Italy). federico.filipponi@isprambiente.it

Abstract

Extended time series of Earth Observation products are increasingly providing consistent information to support applications in a wide variety of domains and disciplines. Despite there is a demanding need for data analytics to extract information from large EO product datasets, few open tools are collecting available techniques to handle and explore raster time series and a small number of scientific studies are comparing results generated from Earth Observation datasets using different data analytics. This research study illustrates and compare the capabilities of different analytics to perform spatio-temporal analysis, included in a newly developed 'rtsa' (Raster Time Series Analysis) package for R programming language .

Results from the analysis of EO multitemporal series using techniques like Empirical Orthogonal Function, Self Organized Maps and Seasonal Trend Decomposition using Loess are presented and compared, providing a picture of the strenghts and weaknesses of the different techniques.

Introduction

Extended time series of Earth Observation products are increasingly providing consistent information to support applications in a wide variety of domains and disciplines. The establishment of the Copernicus European Programme, served by specifically designed Sentinel satellites, activated the generation of large harmonized spatio-temporal datasets, freely available to users under six thematic services for the analysis of spatial features and temporal patterns as well as the monitoring of changes and anomalies. In last decades climate modeling employed many mathematical and statistical methodologies to extract concise information from large spatio-temporal datasets. More recently the availability of extended Earth Observation time series took advantage from the data analytics developed for climate modeling to analyse spatial features and temporal patterns. Despite there is a demanding need for data analytics to extract information from large EO product datasets, few open tools are collecting available techniques to handle raster time series. In addition, many techniques for spatio-temporal analysis can not handle incomplete time series and require appropriate gap-filling methodologies to interpolate raster time series. Capabilities of a newly developed 'rtsa' (Raster Time Series Analysis) package for R programming language providing a collection of analytics to perform spatio-temporal analysis from raster time series is here presented. The package acts as a front-end for already existing R function for gap-filling and multitemporal series analysis, collecting methodologies for gap-filling of incomplete Earth Observation time series (like linear, polynomial and stine interpolation, DINEOF, Gapfill method), analytics for time series analysis using temporal approaches (anomaly detection, Signal Trend Decomposition using Loess, Breaks For Additive Season and Trend BFAST, X-11, X-13ARIMA-SEATS) and spatio-temporal approaches (Empirical Orthogonal Function, Empirical Orthogonal Teleconnections, Self Organizing Maps). Results from the analysis of EO multitemporal series using techniques like Empirical Orthogonal Function, Self Organized Maps and Seasonal Trend Decomposition using Loess are presented and compared, providing a picture of the strenghts and weaknesses of the different techniques.

Materials and Methods

A time series of SST observations has been downloaded from the Copernicus Marine Environment Monitoring Service (CMEMS). The SST refer to CMEMS operational product 'SST_MED_SST_L4_REP_OBSERVATIONS_010_021' reproducing daily gap-free measurements of SST at 4 km horizontal resolution over the Mediterranean Sea, obtained from infra-red measurements collected by numerous satellite radiometers and gap-filled using optimal interpolation method.

Several statistical approaches can be applied in order to extract information from time series. In this research study time series of SST products were analyzed in order to determine spatial and temporal variability using three different analytical methods: 1) Empirical Orthogonal Function (EOF) analysis; 2) a series of Self-Organizing Maps (SOM); 3) Seasonal Trend Decomposition analysis using Loess (STL).

Empirical Orthogonal Function (EOF) analysis (Bjoernsson & Venegas, 1997) has been adopted in order to reduce the input time series dataset to a smaller set of orthogonal patterns. Such analysis is a principal component analysis applied to datasets representing both spatial and temporal dimension, in order to perform a decomposition into dominant spatial-temporal modes. Each principal component loading pattern is named mode, and can be both represented in its spatial dimension and temporal dimension, the latter is denoted expansion coefficient time series or principal component amplitude. This tool has been largely used in climatology, oceanography and EO disciplines to rank spatial patterns of variability, their time variation and the importance of each pattern on the basis of variance. Results from EOF analysis have been scaled between values '-1' and '1' in order to compare the different modes result and facilitate the interpretation of the spatial patterns of variability.

The Self-Organizing Maps (SOM) are a method that can be used to reduce data dimensionality of large spatio-temporal datasets adopting a neural network method with an unsupervised training process (Kohonen, 2001). It generates a two dimensional series of maps that represent the main spatial pattern of variability, maintaining the topological features and, diversely from EOFs, the same units of the original data. Each observation in time from the input time series dataset is associated to the most similar, namely the Best Matching Unit (BMU) recording a weight vector with the smallest distance, among the generated series of maps. SOM analysis requires the user to set the size of the two dimensional grid that will be used to characterize the time series dataset.

Seasonal Trend Decomposition using Loess (STL) (Cleveland et al., 1990) is one of the most widely used method for signal decomposition analysis. It divide up a time series into three components, namely the trend, seasonality and remainder.

Results

1. Empirical Orthogonal Function

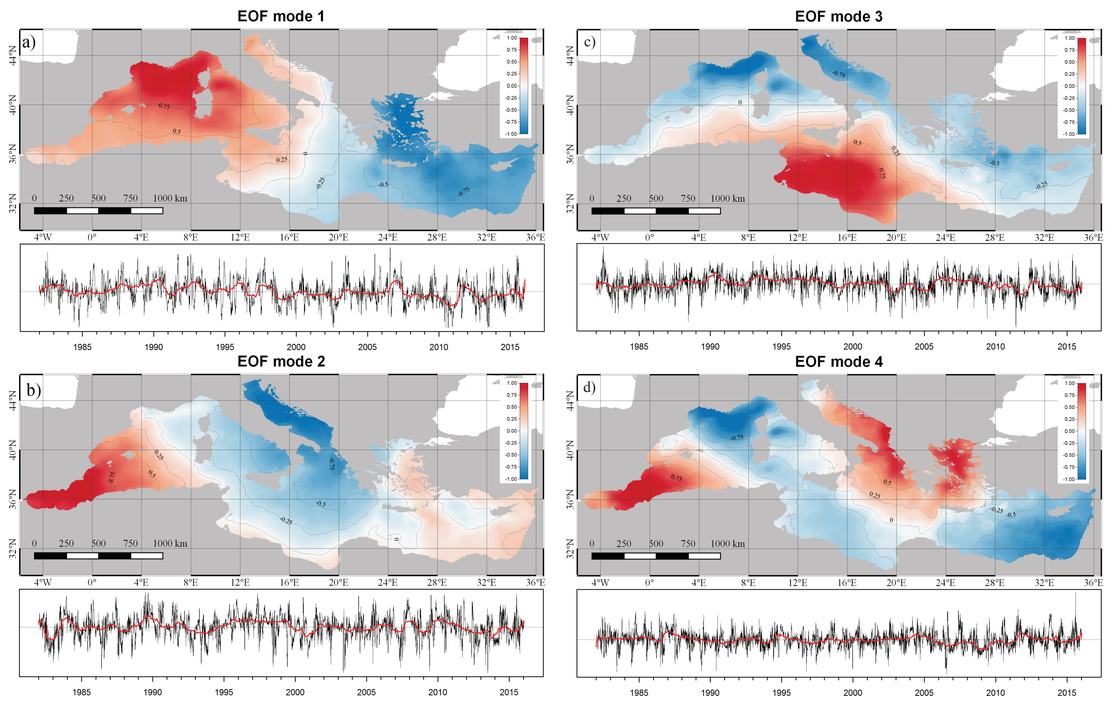


Figure 1. Modes and Expansion Coefficients resulting from EOF analysis: a) EOF mode 1 and the correspondent EC. b) EOF mode 2 and the correspondent EC. c) EOF mode 3 and the correspondent EC. d) EOF mode 4 and the correspondent EC.

2. Self Organizing Maps

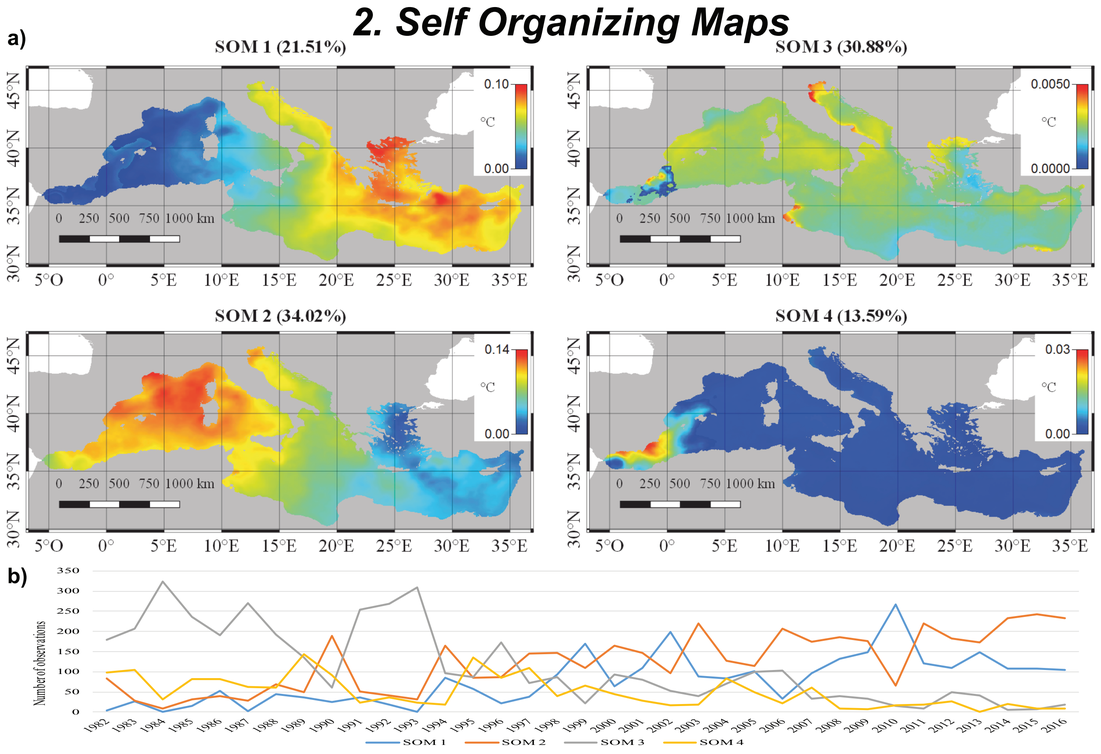
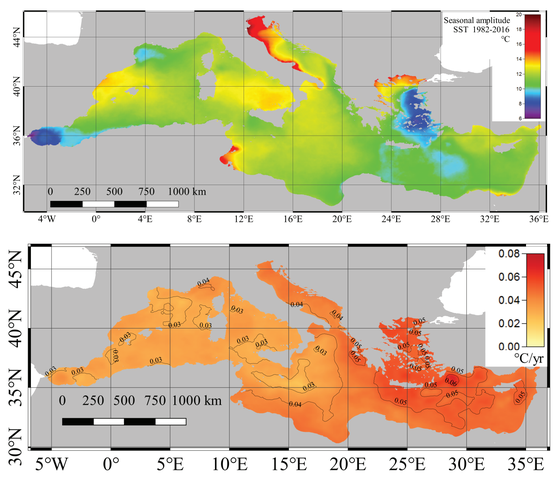


Figure 2. a) Spatial representation of resulting 2x2 SOMs. b) Number of annual observation for each SOM derived from BMU statistics.

Analytical method	Strengths	Weaknesses
Seasonal Trend Decomposition	• Capable of extracting the seasonal, trend and residual component • Valuable for the extraction of the seasonal signal in time series • Allow the extraction of temporal trends • Results are not scaled, in the input values range	• Need gap-free multitemporal time series • Work on single pixel temporal profile, not accounting for spatial variability • User need to define the time windows for the extraction of seasonal and trend signal
Empirical Orthogonal Function	• Capable of extracting the most common patterns • Valuable for reducing dimensionality for big datasets • Allow the interpretation of both spatial and temporal patterns	• Need gap-free multitemporal time series • Compute-intensive (lot of memory) • Lack in finding the least occurring patterns or non-linear patterns • Results are scaled, difficult to link to the original values • Generally does not have the ability to split the contribution of different forcings into the different modes (modes are mixed contributions) • Need result interpretation by expert user to identify relations with environmental forcing since resulting modes do not always represent real and clear patterns
Self Organizing Maps	• Capable of extracting the most common patterns • Valuable for reducing dimensionality for big datasets • Results are not scaled, in the range of the original values • Ability to link single observations to the resulting maps in time dimension	• Need gap-free multitemporal time series • Compute-intensive • User need to initialize the training • Expert knowledge is needed to identify relations with environmental forcings • Resulting temporal information is discrete

Table 1. Strengths and weaknesses of the adopted analytical methods.

3. Seasonal Trend Decomposition using Loess



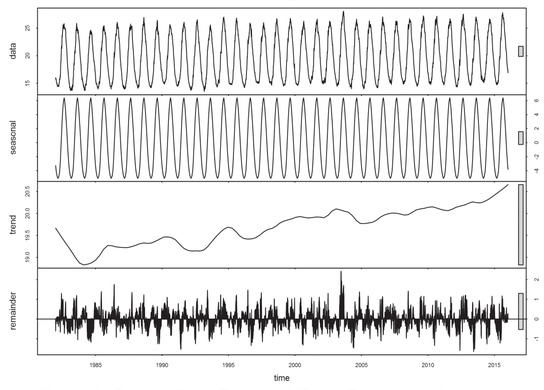


Figure 3. Results from Seasonal Trend Decomposition analysis: a) SST seasonal amplitude for the period 1982-2016. b) SST trend for the period 1982-2016. c) Decomposed temporal profile of daily Mediterranean average SST for the period 1982-2016.

Discussion

EOF modes generated from deseasonalized time series, even representing most of the overall variability, do not show clear trends in the expansion coefficients. EOF mode 1 show a sharp gradient between the western and eastern part of the Mediterranean basin (Fig. 1a). Generally EOF do not have the ability to split the contribution of different forcings into the different modes (modes are mixed contributions). Even if EOF is a linear method capable of extracting the most common patterns, it is somewhat lacking in finding the least occurring or non-linear patterns since it reduces most of the variability to the first few modes and their interpretation can be critical as they will not always represent real and clear patterns in the field modes. Here spatial patterns are strongly related to atmospheric circulation, but comparison with other variables and expert user knowledge is required for a more complete interpretation of the patterns, since they comes from the mixed contributions of different physical forcings, even if EOF modes cannot always be interpreted in terms of real physical signal.

Temporal BMUs of SOMs generated from anomaly SST time series are following the seasonal signal, indicating that the SOM method lacks in finding the least occurring patterns when either the principal signal in time series is not removed or a representation set up with a small number of training samples is adopted. Diversely, SOM generated from deseasonalized SST time series (Fig. 2) show different spatial patterns and temporal trends in the temporal BMUs. SOM 3 strongly related to seasonal amplitude, and SOM 4 detect a least occurring pattern. In addition, SOM method seems to be sensitive to the difference in the source products, especially those acquired from different satellite sensors, even if a bias removal procedure is applied, as for the case of the multi-sensor CMEMS product used in this study.

An overall comparison on the strengths and weaknesses of the different multitemporal analysis methodologies, coming from authors experience gained also in other research studies (Valentini et al., 2016; Filipponi et al., 2017), are reported in Tab. 1. There are still few comparisons over analytical methods for multitemporal analysis as well as demanding needs of data analytics to extract information from large EO products datasets, using open tools dealing with raster time series.

Seasonality signal (Fig. 3a) contains most of the SST time series variability in temporal signal. In order to analyze the temporal trends it is advised to remove from the raster time series the seasonality signal (to obtain deseasonalized time series) rather than the average value (to obtain anomaly time series).

From the deseasonalized SST time series it was possible to identify temporal trends and the main spatial patterns of variability. More specifically, SST in the Mediterranean Sea is rising at a rate of 0.04 degrees×year^{-1} on average during the analyzed period (Fig. 3b). A major increase of temperature is located in the eastern part of the basin.

Conclusion

In this research study the main spatial and temporal patterns of variability have been extracted using different analytical approaches from raster time series. The Mediterranean Sea show a SST rise of 1.4 Celsius degrees on average during the period 1982-2016, with a higher increasing trend in the eastern part of the basin. The used data analytics are capable to extract information from large EO product datasets and their strengths and weaknesses reported, even if the interpretation of results often requires expert user knowledge. This research study demonstrates that the use of extended time series of EO products, and the availability of wide datasets collection in the framework of the Copernicus services, allow the generation of added value product from the multitemporal analysis of EO datasets. The newly developed 'rtsa' (Raster Time Series Analysis) package for R programming language, currently available on GitHub, provides a collection of analytics to perform spatio-temporal analysis and may take part to the new era of big data from space.

References

Bjoernsson H., Venegas S.A., 1997. A manual for EOF and SVD analyses of climate data, McGill University, CCGCR Report No. 97-1, 52pp.

Cleveland R.B., Cleveland W.S., McRae J.E., Terpenning I., 1990. STL: A Seasonal-Trend Decomposition Procedure Based on Loess, Journal of Official Statistics, vol. 6, pp. 3-73.

Filipponi F., Valentini E., Taramelli A., 2017. Sea Surface Temperature changes analysis, an Essential Climate Variable for Ecosystem Services provisioning, Proceedings of "MultiTemp 2017 Conference" (in press).

Kohonen T., 2001. Self-organizing maps, 3rd Ed., New York: Springer.

Valentini E., Filipponi F., Nguyen Xuan A., Taramelli A., 2016. Marine food provision ecosystem services assessment using EO products, Proceedings of "ESA Living Planet Symposium 2016", ESA SP-740 (CD-ROM).

Acknowledgements

The scientific and technical knowledge contained derives from the outcomes of the European Union's Horizon 2020 research and innovation programme under grant agreement No. 641762 'ECOPOTENTIAL - Improving future ecosystem benefits through earth observations' and European Union's founded project 'MERMAID - Innovative Multi-purposE off-shoRe platforMs: plAnnIng, Design and operation', FP7 OCEAN, Grant Agreement No. 288710. The dissemination and the uptake of the products are funded under the Copernicus Training and Information Sessions (Copernicus User Uptake Framework Contracts). This research study has been conducted using the Copernicus Marine Environment Monitoring Service products (CMEMS). Thanks to ISPRA colleagues Emiliana Valentini, Andrea Taramelli and Alessandra Nguyen Xuan for fruitful discussion and support.