Project Title:	ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS
Project number:	641762
Project Acronym:	ECOPOTENTIAL
Proposal full title:	IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS
Type:	Research and innovation actions
Work program topics addressed:	SC5-16-2014: "Making Earth Observation and Monitoring Data usable for ecosystem modelling and services"

Deliverable 4.5

Time series EOFs (map and analysis)

Due date of deliverable:	30th of November 2017
Actual submission date:	12th of December 2017
Version:	V1
Main Authors:	E. Valentini, F. Filipponi, A. Nguyen Xuan (ISPRA); I. Dontas, N. Bogonikolos (ARATOS); I. Manakos, G. Kordelas, K. Marini, V. Kotsias, M. Papanastasiou (CERTH); J. Bustamante, R. Díaz-Delgado (CSIC); A. Karnieli, N. Ohana-Levi, A. Stein (BGU); C. Tarantino, P. Blonda (CNR-IIA)

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762

Project ref. number	**641762**
Project title	**ECOPOTENTIAL: IMPROVING FUTURE ECOSYSTEM BENEFITS THROUGH EARTH OBSERVATIONS**

Deliverable title	*Time series EOFs (map and analysis)*
Deliverable number	D4.5
Deliverable version	V1.0
Contractual date of delivery	30.11.2017
Actual date of delivery	12.12.2017
Document status	Final
Document version	1.0
Online access	ECOPOTENTIAL website: http://www.ecopotential-project.eu
Diffusion	Public
Nature of deliverable	Report
Workpackage	WP4
Partner responsible	Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA)
Author(s)	E. Valentini, F. Filipponi, A. Nguyen Xuan (ISPRA); I. Dontas, N. Bogonikolos (ARATOS); I. Manakos, G. Kordelas, K. Marini, V. Kotsias, M. Papanastasiou (CERTH); J. Bustamante, R. Díaz-Delgado (CSIC); A. Karnieli, N. Ohana-Levi, A. Stein (BGU); C. Tarantino, P. Blonda (CNR-IIA)
Editor	E. Valentini (ISPRA)
Approved by	
EC Project Officer	Gaëlle Le Bouler

Abstract	The present Deliverable, D4.5, illustrates the activity carried out, within Task 4.2 and Task 4.3, in the first 30 months of the project. It provides an overview of the algorithms adopted or developed to analyse time series of continuous environmental variables and demonstrate their applications for selected case study. The techniques collection contributes to the provision of products and tools for the assessment of ecosystem state and changes,

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762

	which play a key role in ecosystem service assessment.
Keywords	Earth Observation, time series, multitemporal analysis, temporal approaches, spatio-temporal approaches, gap-filling

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641762

Table of Contents

Co-funded by the European Union

1. Executive summary

WP4 aims at exploiting existing Earth Observation (EO) data combining them with in situ measurements (WP5) on a large set of Protected Areas (PAs) to use ecosystem-relevant knowledge for modeling and assessing ecosystems status and ecosystem services. This is done by using pre-existing and new techniques and algorithms and by delivering high-level and easy-to-access EO tools and products.

WP4 products - and associated techniques and algorithms are available through the development of a technical exploitation platform named Earth Observation Data for Ecosystem Monitoring (EODESM), which is a pre-operation multimodular and multi data source system, based on open source software, to quantify variables relevant to ecosystems state and change.
As EODESM system should enable ecosystem mapping and monitoring, in order to evaluate status and change in ecosystem and ecosystems services - and observe cause-effect relationships - the understanding of their variation in space and time is a key element in EODESM.
Thus WP4 is linked to (providing basis for) and supports the activities within WP6 (EO based ecosystem modeling), WP7 (ecosystem services), WP8 (cross-scale interaction), WP9 (requirements for future protected areas) and WP10 (Virtual Laboratory Platform). The D4.5 also provides useful information supporting the Protected Area Managers community (WP9) in making effective monitoring, planning and managing (WP11, 12).

Ecosystem and related variables change is addressed in D4.3 for classified maps and some continuous environmental variables, in the form of change of classes and change of variable values, typically done for pairs of observations in time. The deliverable D4.5 integrates D4.3 with time series techniques for the analysis of continuous variables.

Deliverable D4.5 is associated to multiple tasks in the framework of WP4, and it presents the time series analysis techniques and algorithms adopted or developed for assessing continuous terrestrial and marine bio-geophysical variables (environmental variables in D4.2) and other relevant indexes and indicators (see Task 4.2 and Task 4.3) derived from EO.
The described techniques contribute to the provision of products for the assessment of ecosystem conditions (state and changes) which play a key role in ecosystem service assessment (see task 4.3.4).

The combination of Essential Variables (EVs) products with multitemporal added-value indicators are useful input for Ecosystem Service modelling (WP6) and assessment (WP7). Moreover, this combination enables the generation of future scenarios for Ecosystem Service assessment (WP8) as it provides results relevant also to contents of D8.2, which is expected to show past and current changes of ecosystems and ecosystem services delivered. Time series analysis of EVs provides insight also to WP9 (Requirements of future protected areas) by helping the monitoring of changes in climate, environment, biodiversity and society that impact PA in several ways. These changes and the related effects indeed could imply that the requirements and criteria for PA designation may have to be updated.
D4.5 is linked also to WP10 (Virtual Laboratory Platform) as the techniques and algorithms it presents should be among the resources accessible through the platform itself; and to WP11 (EO supported policy development and integration) as the products it shows are among the material for facilitating and enhancing the integration of EO data into policy and decision making processes. Finally D4.5 contents could support capacity building, training and results exploitation activities of WP12 (Capacity building and knowledge exchange).

The deliverable has the following structure:

- Chapter 2 shows in brief the main objectives of the deliverable and the relevance of EO data time series analysis for ecosystem and ecosystem services mapping, assessment and monitoring;
- Chapter 3 addresses some of the available time series analysis techniques and algorithms (divided in temporal approaches and spatio-temporal approaches) in addition to methodology for data gap-filling as several of these techniques and algorithms do not manage incomplete time series. Furthermore it

describes a developed R programming language package containing the software code to run the some of the described techniques;
- Chapter 4 provides time series analysis applied to relevant Ecopotential sites for both terrestrial and marine environment in order to show with practical examples the capability of the different methods reported in Chapter 3, according to the required information in the frame of ecosystem service assessment;
- Chapter 5, finally, identifies for selected variables, which techniques and algorithms could be used in order to obtain the relevant information (e.g. seasonality, overall trend, main temporal patterns, etc.). and shows a comparison on the strengths and weaknesses for some of the different multitemporal analysis methodologies.

2. Introduction

To assess and monitor ecosystems and services and the progress and effectiveness of ongoing conservation policies and management activities (e.g. current PAs) - thus supporting the environmental adaptive management - or to design new policies and plans (e.g. future PAs), monitoring programmes have to be established.
As economic resources for monitoring activities are limited and the monitoring effort should be aimed at providing information useful to several stakeholders across several domains and scales (e.g. PA goals are framed in the national environmental goals), Essential Variables (EVs) are to be used.
Essential Variables, being able to synthesize the information are increasingly adopted in different thematic communities (Biodiversity, Ocean, Climate, etc.), and help facing issues like status and ongoing changes of ecosystem functions and services in varied environmental conditions.

Ecosystem and ecosystem functions and services undergo change in space and time. Spatial change are deepened in D4.3. Temporal changes and spatio-temporal change can be discrete and abrupt or continuous and slow: the former can be detected by performing, for example, analysis between two EO observations in time data. The latter is much more challenging to be detected (e.g. seasonality, long term trends, etc.) as it entails the analysis of a large amount of data related to longer period, e.g. time series of daily or monthly observations over a year or tens of years. Uncertainty related to the gap-filling procedures and multitemporal analysis is in some cases generated from the analytical methodology and associated to the results, in order to support further data analysis in numerical modeling and/or the evaluation of changes of environmental conditions.

2.1 Main objectives

The stated operational objectives of WP4, as listed in the DOW, are:

1. Design and develop a pre-operational multi-modular system named EO Data for Ecosystem Monitoring (EODESM) for Level 1, 2, 3 and 4 Remote Sensing (RS) data to quantify Essential Biodiversity Variables and other EVs and document quality and uncertainty. An advancement of the state of the art using both pre-existing and new techniques will be achieved particularly when dealing with very high spatial resolution (e.g., 3rd Party missions: WorldView-3), temporal resolution (e.g., Sentinel) and spectral resolution (e.g., DLR hyperspectral EnMap).
2. Optimise the selection and use of appropriate EO data, considering sensor types, spatial, temporal and spectral resolution, and Copernicus core products for ecosystem classification, mapping and monitoring.
3. Elaborate open source software and products.
4. Support activities in WP6, WP7, WP8 and WP9 and provide results to WP10. Provide EO product validation, inter-comparison and visualization tools for assessing the quality of data products (e.g., land cover and habitat maps). As output, WP4 will provide comprehensive multi-source, multi-scale and multi-temporal EO products (services) for ecosystem monitoring and will distribute them as an ECOPERNICUS service (WP10).

In this framework, D 4.5 contains analysis developed within the activities carries out in Task 4.2 (Retrieval of bio-geophysical terrestrial and marine variables) and Task 4.3 (Land cover/use ecosystem and habitat classification) and helps to reach objectives 1, 3 and 4.

2.2 From space to time: multitemporal analysis of EVs

In the perspective of providing useful information for policies, planning and management decision-makers (e.g. Protected Area Managers), effective monitoring and assessment of the state and trends of ecosystem conditions and services are highly desirable. Therefore to support the monitoring and assessing of ecosystems and services there is the need to understand the main temporal patterns (occurred in the past), how they change in space (see D.4.3) and how evolve through the time (see this deliverable D4.5).

2.3 List of acronyms

AGF	Asymmetric Gaussian Fitting
AVHRR	Advanced Very High Resolution Radiometer
BFAST	Breaks For Additive Season and Trend
CACAO	Consistent Adjustment of the Climatology to Actual Observations
Chl-a	Chlorophyll – a
CLC	Corine Land Cover
CMEMS	Copernicus Marine Environment Monitoring Service
CMK	Contextual Mann-Kendall
CSSA	Caterpillar Singular Spectrum Analysis
DBEST	Detecting Breakpoints and Estimating Segments in Trend
DINEOF	Data Interpolating Empirical Orthogonal Functions
DOW	Downscale Ocean Waves
EBV	Essential Biodiversity Variable
ECV	Essential Climate Variable
EMD	Empirical Mode Decomposition
EO	Earth Observation System
EOC	Essential Ocean Variable
EODESM	Earth Observation Data for Ecosystem Monitoring
EOF	Empirical Orthogonal Function
EOT	Empirical Orthogonal Teleconnections
EV	Essential Variable
GCOS	Global Climate Observing System
ICSSA	Iterative Caterpillar Singular Spectrum Analysis
ID	Identifier
LAI	Leaf Area Index
NDVI	Normalized Difference Vegetation Index
PA	Protected Area
PROBA-V	Project for On-Board Autonomy - Vegetation
RS	Remote Sensing
SAR	Synthetic Aperture Radar
SGF	Savitzky-Golay Filter
SOM	Self Organized Map
SST	Sea Surface Temperature

STL	Seasonal Trend Decomposition using Loess
SVD	Singular Value Decomposition
TSGF	Temporal Smoothing and Gap-Filling
VIIRS	Visible Infrared Imaging Radiometer Suite
VLP	Virtual Laboratory Platform

3. Methodology for time series analysis of EVs

Here below several techniques and algorithms to perform temporal analysis of data are briefly reviewed. As some of these do not work with no data (having gap is quite common when dealing with EO data), paragraph 3.1 reports gap-filling techniques to be used before the actual temporal or spatio-temporal analysis. Another constraints in the techniques selection is the amount of data.

Paragraph 3.2 reports analytical methodologies for the time series analysis of EVs using temporal approach (Figure 1), using many single datasets, each incorporating the information derived from the temporal profile of a single pixel. Paragraph 3.3 reports analytical methodologies for the spatio-temporal approach (Figure 1), using a single datasets, the so-called 'datacube' incorporating the information derived from the entire spatial and temporal variability, namely the temporal profile of all pixels.

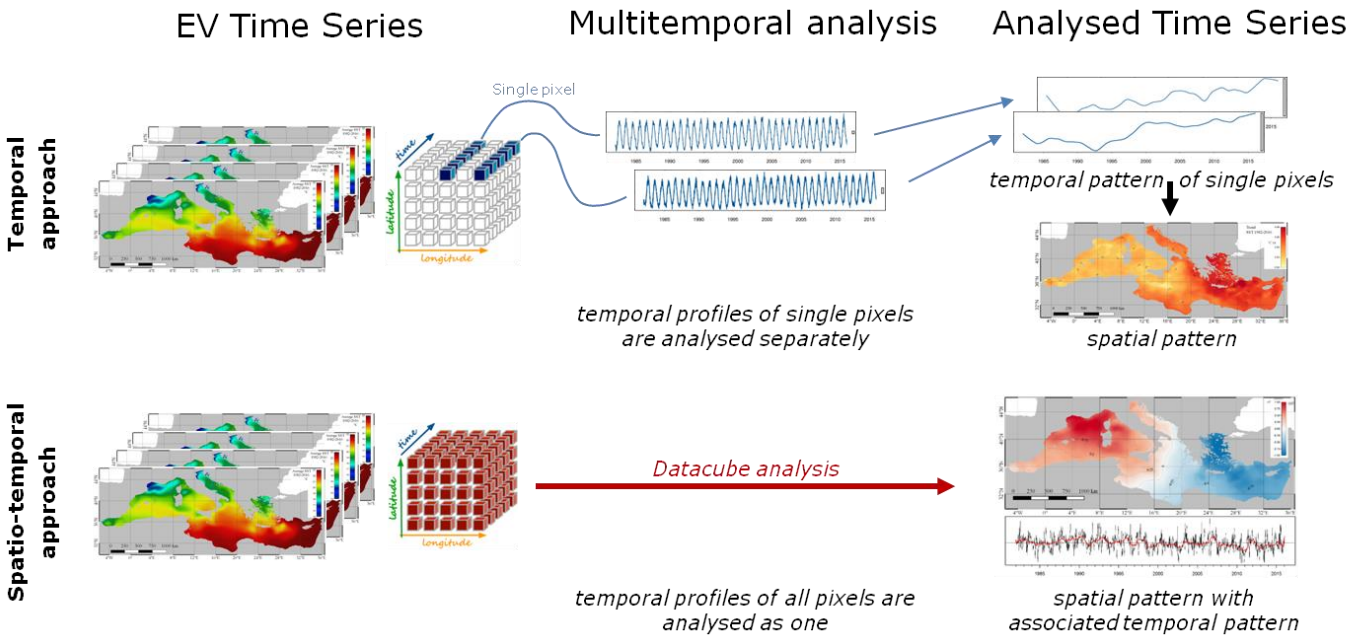

Figure 1. Overview of time series analysis of EVs using temporal approach and spatio-temporal approach.

3.1 Gap-filling

Satellite Earth Observation is a technology used to study a wide range of earth surface processes. It is usually used a long distance from the ground, and when compared to ground based measurements, the technology has the advantage of large spatial and temporal coverage. Time series of satellite data provide valuable information on ecosystem cover, use and condition that could advance understanding of ecosystem function, resilience and dynamics. Time series of different EVs (e.g., Sea surface Temperature, Leaf Area Index) emerge therefore from the systematic monitoring of their temporal evolution. In this way and due to their high temporal resolution, they allow for observing and documenting valuable data about the dynamics of the environment and its changes, offering a valuable instrument to the PA managers and the work dedicated to the protection of the environment.

Actually EO time series are often plagued with a percentage of missing values (caused for example by cloud cover)

creating sporadic and/or continuous gaps in their regular space time-step. Many practical applications (e.g., extreme value analysis) as well as statistical methodologies (e.g., spectral analysis, calibration-learning algorithms, stochastic modeling and downscaling) have no tolerance to missing values. Processing of EVs datasets by gap-filling their missing values is thus a necessary procedure. One of the main issues about the development of gap-filling methodologies is related to the size of the dataset and consequently the usability of the approaches due to large computational time and wide memory resources availability. Generally spatial resolution of EVs over marine areas is significantly lower than the spatial resolution over terrestrial areas, making the gap-filling methodologies development more challenging for terrestrial EVs.

Several methods have been proposed for gap-filling environmental datasets ranging from rather simple to extremely complex approaches e.g., linear or logistic regression, polynomial or spline interpolation, inverse distance weighting, ordinary kriging, and stochastic models that are fitted to the available records.

Additional statistical techniques arising in the last decade, include artificial neural networks and nearest neighbour techniques or even hybrid methods, employing both process based and statistical tools.

In this section the principles of application of some relevant gap-filling methodologies used in Earth Observation are described.

3.1.1 Temporal gap-filling approaches: linear, spline and stine interpolation

Temporal gap-filling approaches rely on predicting missing values for each single pixel by interpolating the available values observed at different times.

Among temporal gap-filling approaches, linear interpolation is a method of curve fitting using linear polynomials to predict missing values in time series. Given two points in one-dimensional space (time dimension), the linear interpolant is the straight line between these points.

The spline interpolation is another one-dimensional gap-filling method where the interpolant is a special type of piecewise polynomial called a spline. Such approach uses elastic rulers to interpolate missing points with polynomials (Forsythe et al., 1977).

Another one-dimensional interpolation method, suggested by Stineman (1980) and also known as 'stine', is based on polynomials and it results in lower slopes near abrupt steps or spikes in the time series, and therefore a smaller tendency for overshooting. The method leads to much less tendency for spurious oscillations than traditional interpolation methods based on polynomials, such as splines.

3.1.2 Asymmetric Gaussian Fitting (AGF)

This method has been proposed within the TIMESAT toolbox (Jonsson and EKlundh, 2004) and has the capability of handle small gaps. It has been designed to properly work with vegetation related EVs, and its principles are based on the phenological features described by vegetation time series. A Gaussian function is adjusted locally over the growing and senescing parts of each season. The functions are finally merged to get a smooth transition from one season to another (Jonsson and EKlundh, 2002). The original TIMESAT implementation includes three conditions preventing the near-constant and noisy data from being processed. Two of these conditions (minimum seasonality in the data and maximum fraction of missing data of 25 %) are not considered to enlarge its domain of validity in case of missing data, allowing more rigorous comparison with the other methods.

3.1.3 Savitzky-Golay Filter based methods

The Adaptive Savitzky-Golay Filter iteratively applies the Savitzky-Golay filter (Savitzky and Golay, 1964) to match the upper envelope of the time series. The original Savitzky-Golay filter consists in a local polynomial fitting with two parameters: the length of the temporal window used and the order of the polynomial. The adaptive Savitzky-

Golay Filter (Chen et al., 2004) optimize by linear interpolation the values of these parameters for each case to get the best match between observations and reconstructed values. This specific adaptation was designed to minimize the effects of cloud and snow contamination that generally decreases the estimates of EVs describing slow variability, like the vegetation indices such as NDVI as well as biophysical variables such as LAI.

The Temporal Smoothing and Gap-Filling (TSGF) method (Verger et al., 2011) is another adaptation of the Savitzky-Golay filter designed to better handle time series with missing observations, where the polynomial degree was fixed to 2 but the temporal window may be asymmetric and variable in length. The method gives better results if the temporal window is adjusted suitably to include a good number of observations per period (before and after each nominal date) in order to have a better application of the polynomial fitting.

3.1.4 Consistent Adjustment of the Climatology to Actual Observations (CACAO)

A novel climatology fitting approach called Consistent Adjustment of the Climatology to Actual Observations (CACAO) (Verger et al., 2013) is found capable to reduce noise and fill gaps in time series by scaling and shifting the seasonal climatological patterns to the actual observations.

The CACAO method is adopted for smoothing, gap-filling satellite time series of vegetation EVs.

CACAO is based on the fitting of a vegetation phenology model. This model is specific to each pixel and is derived from the climatology computed over the time series of the considered pixel. CACAO method appears to stand between very flexible methods such as the adaptive SG filter and methods based on a unique phenology model such as the AGF method. In terms of the required computer resources, CACAO is as demanding as AGF, although SG is twice faster.

The CACAO method was first applied to simulated time series of daily LAI estimates as derived from AVHRR observations. That simulation experiment showed a significant improvement in the theoretical performances of LAI reconstructions and phenology extraction for CACAO as compared to SG and AGF methods in all the situations, particularly for the higher levels of the fraction of missing data and daily LAI uncertainties.

CACAO was assessed first over simulated daily Leaf Area Index (LAI) time series with varying fractions of missing data and noise. Comparisons with two widely used temporal filtering methods-the Asymmetric Gaussian fitting (AGF) (Jonsson and Eklundh, 2002) and the Savitzky-Golay (SGF) filter (Savitzky and Golay, 1964; Chen et al., 2004) -revealed that CACAO achieved better performances for smoothing AVHRR time series characterized by high level of noise and frequent missing observations.

The main limitation of CACAO reconstruction method is found to be its inability to capture underlying atypical modes of seasonality including rapid natural and human induced disturbances in the LAI time series that strongly differ from the average climatology (e.g., flood or fire events, changes in the land cover). To prevent from such drawback, the resulting fitted climatology can be fused with a product closer to the actual LAI observations as the one resulting from the adaptive SG filtering.

3.1.5 Iterative Caterpillar Singular Spectrum Analysis method (ICSSA)

The Iterative Caterpillar Singular Spectrum Analysis Method (ICSSA) (Kandamasy et al., 2013) is a modification of the Caterpillar Singular Spectrum Analysis (CSSA) method (Golyandina and Osipov, 2007) developed to describe time series and fill missing data by decomposing the time series into Empirical Orthogonal Function modes. Empirical Orthogonal Function is a principal component analysis applied to datasets representing both spatial and temporal dimension, and perform a decomposition into dominant spatial-temporal modes in order to reduce the input time series dataset to a smaller set of orthogonal patterns (see paragraph 3.3.1). It allows filling gaps and forecasting data at the extremities of the time series. This modified version was proposed to correct for overestimation of small drops and better fitting to the peaks as compared to the original CSSA formulation. The method requires 2 parameters: the window length and the number of eigenvectors (orthogonal functions) used for

the reconstruction. Better reconstruction can be obtained for large number of eigenvectors but at the cost of a decrease in the smoothness.

3.1.6 Empirical Mode Decomposition (EMD)

The Empirical Mode Decomposition (EMD) method (Huang et al., 1998) decomposes the time series into a small number of Intrinsic Mode Frequencies (IMFs) derived directly from the time series itself, using an adaptive iterative process where the data are represented by intrinsic mode functions, to which the Hilbert transform can be applied. The method requires the setting of two parameters: the threshold for convergence and the maximum number of IMFs. We should note that the EMD method requires the time series to be continuous.

3.1.7 Data Interpolating Empirical Orthogonal Functions (DINEOF)

Data Interpolating Empirical Orthogonal Functions (DINEOF) (Alvera-Azcarate et al., 2006; Beckers and Rixen, 2003) is an EOF-based technique developed to reconstruct missing data in geophysical datasets. It exploits the spatio-temporal coherency of the data to infer a value at the missing location.

DINEOF is a methodology widely used in oceanography, specifically designed to reproduce missing data in time series of EVs such as SST (Beckers et al., 2006) and Chl-a (Wang and Liu, 2014), since those datasets are generally hampered by many missing pixels due to various reasons, mainly from cloud cover.

The Empirical Orthogonal Function (EOF) analysis is a method to determine a set of orthogonal functions, named modes, that characterizes the co-variability of time series for a set of grid points. It is often used to study possible spatial modes (patterns) of variability and how they change with time.

The DINEOF methodology is based on the fact that an EOF analysis aims to extract a small number of significant modes, present in the physical system, from a large dataset. These reduced modes should represent a large fraction of the original variability of the dataset. The combination of the dominant EOF modes and their amplitudes can therefore help recover missing data values.

DINEOF is a procedure that fills gaps by iteratively decomposing the data field via Singular Value Decomposition (SVD) until a best solution is found as compared to a subset of reference values (non-gappy). This is done by progressively including more EOFs in the reconstruction of the “gappy” locations until minimization of error converges.

Recently it was found that the ‘irlba’ package (‘Implicitly Restarted Lanczos Bidiagonalization Algorithm’ for fast truncated singular value decomposition and principal components analysis for large dense and sparse matrices) for R language can be used instead of SVD. IRLBA package finds a few approximate largest (or, optionally, smallest) singular values and corresponding singular vectors of a sparse or dense matrix using a method for Augmented Implicitly Restarted Lanczos Bidiagonalization (Baglama and Reichel, 2005). It is a fast and memory-efficient way to compute a partial SVD.

Regarding the use of similar methodologies to reproduce missing data for terrestrial EVs, DINEOF method itself was found insufficient, so similar methodologies have been developed.

DINEOF was initially developed in oceanography for the reconstruction of single EV, or multiple EVs in multi-variate cases (Alvera-Azcarate et al., 2009; Beckers et al., 2014). It was found that the original DINEOF method is not ready to reproduce missing data for terrestrial EVs for two main reasons. The first problem is its heavy computational cost. Unlike the oceanographic EVs for which DINEOF was originally designed, land EVs usually have much higher spatial resolutions. It seems inefficient and impractical to manipulate very large matrices at one time with this method. The second problem is that multiple products of the same EV usually have different temporal resolutions. Thus, without temporal interpolation, it is rather difficult to form all of the products into a single matrix.

For these reasons, Hierarchical Empirical Orthogonal Functions (HEOFs) were proposed and developed to deal with

both temporal and spatial patterns. The HEOFs work on two levels: coarse and fine-resolution. The original HEOF procedure can thus be simplified by dividing the dataset into small subsets and applying an EOF to each of them. Also in this case, the dataset at coarser resolution data can require considerable memory capacity. Information derived from different EOF analysis cannot be directly integrated, since there will be inconsistent edges among different sub-datasets. An HEOF-based approach (Wang and Liang, 2011) was proposed to overcome these problems. The traditional EOF analysis is applied at two levels: to the coarse-resolution aggregated data and to the multiple fine-resolution sub-datasets, both of which have a small dimension that is feasible for EOF analysis. The data are first aggregated to coarse resolution, and the EOF analysis is then carried out on the coarse-resolution data. Poor quality blocks are replaced during the EOF procedure. The improved coarse-resolution data are used as the mean to center the fine-resolution subsets. For the fine-resolution subsets, two adjacent blocks are intentionally overlapped. After the EOF analysis is applied, the weighted averages of the overlapping area are calculated as the final filtered values to reduce the 'blocky effect', and each pixel weight is given by its distance to the edge.

3.1.8 Gapfill procedure

The gapfill procedure (Gerber et al., 2016) is a method to reconstruct a complete time series using both spatial and temporal information in the gap-filling process, similarly to other proposed methods (Poggio et al., 2012; Weiss et al. 2014; de Oliveira & Epiphanio, 2012; de Oliveira et al., 2014). The approach uses spatio-temporal subsets around missing values to predict them. A ranking on each temporal observation is used to imitate the seasonal evolution of the spatial field. The ranked images are used as a non-parametric seasonality adjustment, and the value of missing pixels is predicted using a linear quantile-regression. The quantile is estimated by evaluating the empirical cumulative distribution function at each value and averaging these results. The algorithm increases the spatial subset if a minimum number of valid pixel needed for the reconstruction is not met, thus adapting to the local distribution of missing values. The procedure also computes the estimation of uncertainties by calculating the quantile differences among different initial sizes of the input spatial subset, and by calculating the 90% prediction interval based on bootstrap methods.

3.2 Time series analysis of EVs: temporal approaches

3.2.1 Anomaly detection

In the natural sciences an anomaly is the deviation in a quantity from its expected value, like the difference between a measurement and a mean or a model prediction (Wilks, 1995).

Dealing with raster time series, the anomaly is generally computed on a pixel base over the temporal dimension and represents the difference of each value in the time series and the mean of the overall time series. In some cases, especially when dealing with time series with a strong harmonic or seasonal signal, the anomaly can be computed as the difference of each value in the time series and the corresponding value in time in the seasonal signal, computed from the overall time series.

Anomaly is often used as the basis for data transformation, namely a function used to transform the data that is generally invertible and continuous, which is performed to convert data before the use of other data analytics tools.

3.2.2 Seasonal Trend Decomposition using Loess (STL)

One of the most widely adopted signal decomposition analysis is the Seasonal Trend Decomposition using Loess (STL) (Cleveland et al., 1990). It is an algorithm that was developed to divide up a time series into three components, namely the trend, seasonality and remainder using local regression techniques.

STL is a backing-fitting algorithm using an iterative estimation of trend, seasonality and remainder . It consists in a

sequence of smoothing operations employing the locally-weighted regression smoother, or loess (Cleveland and Grosse, 1990). STL consist of two recursive procedures: a inner loop, consisting of a seasonal and a trend smoothing updating the input time series, nested inside a outer loop, computing the robustness weights using iterated weighted least-squares (Cleveland et al., 1990).

More recent enhanced versions of the STL algorithm (Hafen, 2010) can deal with missing values in time series and allow local quadratic smoothing, post-trend smoothing and endpoint blending.

3.2.3 Breaks For Additive Season and Trend (BFAST)

Breaks For Additive Seasonal and Trend (BFAST) is a generic change detection approach (Verbesselt et al., 2010) to iteratively estimate the time and the number of changes from a time series, characterizing their magnitude and direction.

Changes in the trend component can be defined as abrupt when a significant variation occur in a small temporal interval, and can often indicate disturbances such as deforestation, fires, floods, urbanization.

BFAST integrates the iterative decomposition of time series into trend, seasonal and residual components (without the need to select a reference period, set a threshold, or define a change trajectory as for other change detection methods) to detect multiple abrupt changes in the seasonal and trend components and to characterize their occurrence time, magnitude, and direction within the trend component.

BFAST is not specific to a particular data type and can be easily applied to EO time series, even if mostly recommended to analyse vegetation time series.

An alternative method to detect and characterize abrupt changes in time series is the Detecting Breakpoints and Estimating Segments in Trend (DBEST) approach (Jamali et al., 2015). Compared to the BFAST approach, it offers higher control over the level of trend generalization or over the change magnitude that is detected and can have better performances in time series where changes of different types and duration may exists (Jamali et al., 2015), but without considerations for seasonal variations as in BFAST (DeVries et al., 2016).

3.2.4 X-11

X-11 is a backing-fitting algorithm which utilizes semi-parametric modeling and robust estimation to compute trend, seasonality and remainder using seasonal adjustment (Shiskin et al., 1967). It allows to consider the seasonal adjustment of time series, in order to find a trend not influenced by yearly anomalies.

3.2.5 X-13ARIMA-SEATS

The X-13ARIMA-SEATS seasonal adjustment program is an enhanced version of the X-11 Variant of the Census Method II seasonal adjustment program (Shiskin et al., 1967). It is a combination of two adjustment models, the X12-ARIMA (Findley et al., 1998; Dagum, 1978) and TRAMO/SEATS (Gómez and Maravall, 1996), and represent a modified version of the X-11 algorithm. It overcomes the limits of the X-11 method related to the use of a set of weights only related to the considered time lag, and hence having a local character, and the asymmetric sets of weights for the smoothing of first and last observations (Dagum, 1978).

It perform a pre-adjustment of the time series before the seasonal adjustment by generating a statistical model for the whole length of the series. The model uses the regARIMA (AutoRegressive Integrated Moving Average) that models the entire time series and extends it with forecasts and backcasts in order to improve the seasonal adjustment of first and last observations. Finally, X-11 algorithm is used to compute trend, seasonality and remainder using seasonal adjustment. X-13ARIMA-SEATS allows also to evaluate the quality of seasonal adjustment.

3.2.6 Non-parametric Mann-Kendall test

The purpose of the Mann-Kendall (MK) test (Mann 1945, Kendall 1975, Gilbert 1987) is to statistically assess if there is a monotonic upward or downward trend of the variable of interest over time. A monotonic upward or downward trend means that the variable consistently increases or decreases through time, but the trend may or may not be linear. Being the MK a non-parametric (distribution-free) test, the residuals from the fitted regression line should not be necessarily normally distributed, and it has the property of being invariant to (monotonic) power transformations. The MK test underlie the assumption that the observations are identically distributed and not serially correlated over time when no trend is present. If a trend is present, this should not be necessarily linear.

3.3 Time series analysis of EVs: spatio-temporal approaches

3.3.1 Empirical Orthogonal Function (EOF)

Empirical Orthogonal Function (EOF) analysis (Bjoernsson and Venegas, 1997) is a principal component analysis applied to datasets representing both spatial and temporal dimension, and perform a decomposition into dominant spatial-temporal modes in order to reduce the input time series dataset to a smaller set of orthogonal patterns. 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 (Falcieri et al., 2014). Anyway, while the 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 (Bjoernsson and Venegas, 1997, Hopkins et al., 2013).

3.3.2 Empirical Orthogonal Teleconnections (EOT)

Empirical Orthogonal Teleconnections (EOT) analysis utilizes a regression based approach to decompose spatio-temporal fields into a set of independent orthogonal patterns. They are quite similar to EOF, with EOT producing less abstract results. In contrast to EOF, which are orthogonal in both space and time, EOT method is rooted in multiple linear regression and yields solutions that are orthogonal in one direction, either space or time, providing a priori selection of number of output modes.

First the base point (predictor) which explains the most of the variance at all other points (predictands) is found in space dimension by linear regression. The first spatial pattern is the regression coefficient between the base point and all other points, and the first time series is taken to be the time series of the raw data at the base point. The original dataset is next reduced by subtracting out the first estimated mode (Van den Dool et al., 2000). This procedure is repeated until a predefined number of EOTs modes.

3.3.3 Self Organizing Maps (SOM)

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.

3.3.4 Contextual Mann-Kendall (CMK) significant test for spatio-temporal time-series analysis

The Contextual Mann-Kendall (CMK) is an enhanced non-parametric Mann-Kendall test accounting for the spatial variability. It assumes that conditions at a certain location that experienced a trend over time will be similar to neighbouring locations, thereby allowing for spatial contextual information to apply (Neeti and Eastman, 2011). CMK relies on the Theil-Sen (TS) slope estimator, which is the median of the slopes calculated between observation values at all pair-wise time steps, with a total of n(n-1)/2 slopes (Sen, 1968). In CMK, the data are ranked according to time and each point is treated as the reference for the data points in successive time periods. It is a non-parametric test, and robust against outliers, and is suitable for quantifying non-linear phenomena. To avoid biases resulting in specific trends in isolated pixels, the contextual form of the Mann-Kendall test is used, taking into account spatial autocorrelation as a line of evidence and include geographical contextual information. The input for this test is composed of a set of raster images (acquired from space/airborne sensors) for different months/years. The data are not required to be ordered with equal temporal intervals; however the dataset must be chronologically ranked and the spatial dimensions of all images, including resolution and coordinates, must identical. Each pixel is then evaluated according to a 3 by 3 neighbourhood. The resulting statistic is then computed for a Z score test statistic that followed a normal distribution with a mean of 0 (Neeti and Eastman, 2011). A pixel with a positive Z score indicates a positive trend whereas negative scores suggest a negative trend. The p-value of the CMK is determined using the normal cumulative distribution function. The resulting map provides a Z score for each pixel, indicating the magnitude and direction of the trend through time for each pixel, with respect to its neighbouring data. This procedure may be conducted in IDRISI's TerrSet 18.3, using Earth Trend Modeler (Eastman, 2016).

3.3.5 Time-series analysis approach for relating Spectral Vegetation Indices (VIs) to Climate Data

Time series of VIs (e.g. NDVI, NDWI) reflect the health of the ecosystem and can be used to derive the timing of phenological events. Climatic variables are the driving factors of phenological events and are closely related to ecosystem health, function, and structure. Time series analysis approach can be applied to understand and model the relations between VIs and different climatic variables (e.g. temperature, precipitation, radiation). Such models are valuable for understanding the role of different climatic variables in different ecosystems, for predicting phenological events, and for comparison of observed time series to the predicted one in case of disturbances (fires, droughts, heat waves). The statistical time series approach is applied according to the following stages: (1) Dividing the time series data to training and validation periods to assess the predictive skills of the model during the validation period (2) Studying the relations between the VIs time series and different prior moving average (Δ) and time lags of climatic variables (τ). This procedure identifies the most influenceable climatic variables while considering their accumulative effect and possible time lag effect. (3) Identifying the dominant periodic signal in the time series of the VIs and the climate variables using the Fast Fourier Transform (FFT), which decomposes the time series into a sum of harmonic signals with different frequencies. Subsequently, anomaly of climatic variables from their dominant cycles is computed. (4) Examining the effect annual temporal trends on VIs using third order polynomial of the time index. (5) The final VIs statistical time series model is composed of the periodic component, the anomalies of the climate variables, averaged over a period Δ and lagged by time τ, such that their cross-correlation with the VIs is the largest, and, in cases of significance, also an annual time polynomial. Stages 2-4 are applied on the training data, while the predictive performance of the model is evaluated during the validation period.

3.4 Conceptual model

Following the description of analytical techniques for multitemporal analysis of continuous EVs, we link their use to the extraction of relevant information for Ecosystem Services assessment. We provide a comparison of the different multitemporal analysis methodologies suitability to derive spatial and temporal descriptors of selected EVs related to terrestrial and marine environments.

Table 1 reports the suitability of different algorithms (available in 'rtsa' package described in paragraph 3.5) to extract required information for the assessment of the ecosystem condition and ecosystem service for the variables SST and NDVI. The table should provide an easier overview to the users of the techniques that could be likely used to retrieve specific information, and can be further improved with other EVs in the forthcoming project phases.

					Temporal dimension						
Ecosystem	**Ecosystem service**	**Ecosystem condition**	**EV**	**EV as EO data**	**Temporal trends**	**Seasonality**	**Main temporal signals**	**Temporal anomalies**	**Abrupt changes**	**Main spatial patterns**	**Spatial anomalies**
Marine	Biomass provision	Thermal condition	Temperature	SST	● ● ● ●	● ● ● ●	● ● ●	● ● ● ●		● ● ●	● ● ●
Terrestrial	Biomass provision	Primary production	Vegetation	NDVI	● ● ● ●	● ● ● ●	● ● ●	● ● ● ●	●	● ● ●	● ● ●

Table 1. Overview of techniques vs Essential Variables vs required information.

Legend
Anomaly
STL
BFAST
X-11
X-13A-S
EOF
EOT
SOM

3.5 R language package

The 'rtsa' (Raster Time Series Analysis) package for R programming language is a collection of analytics to perform spatio-temporal analysis from raster time series and freely available for R software language.

The package 'rtsa' acts as a front-end to already available functions in various R packages, specifically designed to handle geographic datasets provided as raster time series. The available functions within the package allow the direct input of raster time series to extract concise and comprehensive information taking advantages of many analytical techniques.

The achievement of the ISPRA team over the time series analysis is focused on the management of geographic datasets provided as raster time series to extract comprehensive information for Ecosystem Services assessment.

Since none of the analytical techniques described in the present document and included in the R package 'rtsa' has been developed by Ecopotential partners, the TRL cannot be declared for each algorithm. The TRL related to the functionality of the 'rtsa' package is declared instead, in terms of front-end capabilities.

Main features of the package are:

- use of raster time series with explicit temporal dimension;
- use of raster time series as direct input to functions;
- use of a raster mask to select the region of interest and reduce memory loads;

- parallel processing using multiple CPUs.

In order to support the use of gappy EO datasets, the package implements a gap-filling procedure using the following methods:

- DINEOF;
- linear interpolation;
- spline interpolation;
- stine interpolation.

The following analytical methods are currently supported:

- Empirical Orthogonal Function;
- Empirical Orthogonal Teleconnections;
- Self Organizing Maps;
- Seasonal Trend Decomposition using Loess;
- X-11;
- X-13ARIMA-SEATS seasonal adjustment.

The package should serve as a EO Data processing tools (in the form of a Cloud Sandbox, see WP3) to enable the use of EO products for the analysis of Ecosystem Services inside and beyond the PAs, and to deal with data intensive requirements. R software language and essential packages are available and supported in the cloud platform, as stated in D3.2.

According to the requirements of the ECOPOTENTIAL Virtual Laboratory (WP10), the software packages available in the GitHub should follow the ECOPOTENTIAL convention, that will be specified in the coming project phases.

The package is freely available for download and installation on the GitHub repository:

https://github.com/ec-ecopotential/rtsa

- Technological readiness level: 5
- Open source (Y/N): Y
- Commercial or proprietary (C/P): P
- Background or Foreground knowledge (B/F): B
- Partners: ISPRA

4. Time series analysis within ECOPOTENTIAL sites

4.1 Terrestrial

4.1.1 Monitoring the reaction of the biomes to external factors: BFAST for the detection of abrupt trend changes in NDVI time series in the case of Doñana marshes

The Doñana National Park in Spain, is a protected area since 1969, a UNESCO Biosphere Reserve since 1980, a Ramsar Site since 1982, a Natural World Heritage Site since 1984 and it is integrated in the Natura 2000 network,

which also includes the surrounding Natural Park with a similar extent.

It contains one of the largest wetland in Western Europe (García and Marín 2005), an intricate matrix of marshes (270 km^2), phreatic lagoons, and a 25 km-long coastal dune ecosystem with its shoreline and representative Mediterranean terrestrial plant communities (around 100 km^2; Diaz-Delgado 2010).

Biodiversity in general, and waterbird diversity in particular, was the main reason to protect Doñana wetlands, that is also the main reason behind Doñana's attractiveness for ecotourism. Such biodiversity is critically dependent on the high wetland productivity (primary and secondary), as well as the high spatial and temporal diversity of wetland habitats. Different waterbird species use habitats with different characteristics (water depth, vegetation cover, salinity, duration of flooding period, etc.) and also at different times of the year (breeding, stop-over, wintering). Moreover, the high inter-annual variability in flooding regime is also important to maintain the diversity of the waterbird community.

Global (climate change), regional (water extraction, eutrophication) and local (modification of hydrological and grazing regimes) pressures could act in synergy and push the ecosystem to undesirable states. Park management can act on local stressors (water management, livestock density, control of alien species) while regional authorities and policies may act on regional ones (water extraction, urban waste water treatment, pesticide/fertilizer regulation).

Monitoring of changes among annual vegetation cycles of consequent years in Protected Areas is valuable for the recognition of patterns, which represent the reaction of the biomes to external factors, such as changes in the meteorological conditions (e.g. the precipitation regime), human intervention or extreme events (e.g. fire). It is an indicator of the primary production of the area and other relevant functions of the ecosystem. The BFAST, Breaks For Additive Seasonal and Trend, approach can be used for monitoring changes, since it is globally applicable and able to analyze each pixel individually without the need to set thresholds for detecting changes within time series. Thus, BFAST is applied for the detection of abrupt trend changes in NDVI time series in the case of Doñana marshes, as a proxy to phenological metrics per pixel. This process includes following steps:

1. **Generation of NDVI time series:** NDVI time series are calculated based on Landsat 5, 7, 8 TM, ETM and OLI sensors' acquisitions. They cover the last decade, namely the period from the beginning of 2007 to the end of 2016. These data are cropped to the limits of a rectangle containing the Doñana marshes area, utilizing Google Earth Engine platform. Images attenuated by clouds are excluded from the NDVI series. The presence of clouds is inferred by taking into account the cloud cover percentage mentioned in the metadata information accompanying each Landsat product, followed by visual inspection. The dates, for which NDVI data is available are given in Table 2.

Year	Date in DD/MM format													
2007	10/03	05/05	29/05	22/06	30/06	08/07	24/07	09/08	20/10	05/11				
2008	09/02	13/04	18/07	03/08	19/08	09/04	06/10	23/11						
2009	11/02	31/03	02/05	18/05	11/06	27/06	13/07	29/07	30/08	23/09	02/10	02/11		
2010	29/01	05/05	21/05	06/06	30/06	16/07	01/08	25/08	10/09	05/11				
2011	09/02	25/02	21/03	06/04	08/05	24/05	04/08	28/08	05/09	13/09	15/10	31/10		
2012	03/01	19/01	04/02	20/02	07/03	24/04	14/08	15/09	01/10	17/10				
2013	19/04	05/05	22/06	08/07	24/07	09/08	10/09	28/10	13/11	29/11	15/12	31/12		
2014	06/04	08/05	11/07	27/07	12/08	28/08	29/09	02/12						
2015	03/01	19/01	04/02	08/03	11/05	27/05	28/06	14/07	30/07	15/08	02/10	19/11	05/12	21/12
2016	23/02	10/03	14/06	30/06	16/07	01/08	17/08	02/09	04/10	23/12				

Table 2. Dates of cloud free NDVI data from Landsat 5, 7, 8 TM, ETM and OLI sensors.

2. **Preparation of NDVI time series for the BFAST method implemented in R:** NDVI time series, prior their input to the BFAST algorithm, should be interpolated in the following way: input data samples are generated per pixel, where each sample contains the date and its corresponding NDVI value. These data samples are used to fit a curve relying on smoothing splines. The resulting data curve, which is continuous, is sampled per 16 days to derive a vector of 228 NDVI values covering the period from 01/01/2007 to 23/12/2016.

3. **Estimation of abrupt changes in NDVI time series:** BFAST approach estimates per pixel's NDVI time series, the IDs of the samples, where abrupt changes in the NDVI time series occur, and the ID of the sample for which the biggest change in the trend component is detected. A sample ID, representing multiple pixels for which abrupt changes are detected, has the form "Year, Day from beginning of Year" and is calculated by:

   $\left(floor\left(2007+16\cdot \mathrm{ID}/365\right), \left(\left(2007+16\cdot \mathrm{ID}/365\right) - floor\left(2007+16\cdot \mathrm{ID}/365\right)\right)\cdot 365\right)$, where

   $floor(x)$ rounds x to the nearest integer less than or equal to x.

   In particular, BFAST method iteratively decomposes the NDVI time series into trend, seasonal and remainder components and estimates abrupt trend changes in the trend component. Figure 2 shows the decomposition of the NDVI time series corresponding to a sample pixel.
4. **Output data:** BFAST outputs are used to generate: (i) a raster containing the total number of detected abrupt changes per pixel (see Figure 3(a)), (ii) a raster containing the time for which the biggest change is detected per pixel (see Figure 3 (b)), where each time value has 7 integers with the first 4 digits representing the year and the last 3 digits demonstrating the day of the year where the change occurs, (iii) a histogram comprising the total number of pixels per sample ID for which abrupt changes are detected (see Figure 5), (iv) a histogram comprising the total number of pixels per sample ID for which the biggest changes are detected (see Figure 6).

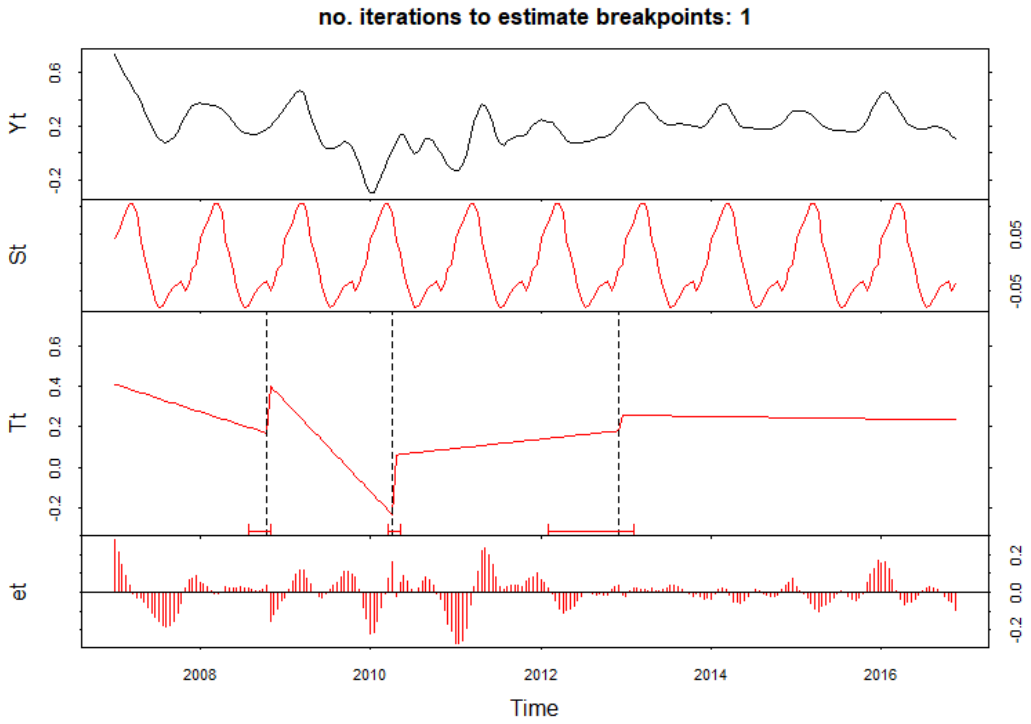

Figure 2. NDVI time series (Yt) for a pixel is shown in black. The estimated seasonal (St), trend (Tt) and remainder (et) components are shown in red. The time of estimated trend changes is indicated with vertical dotted lines.

Co-funded by the European Union

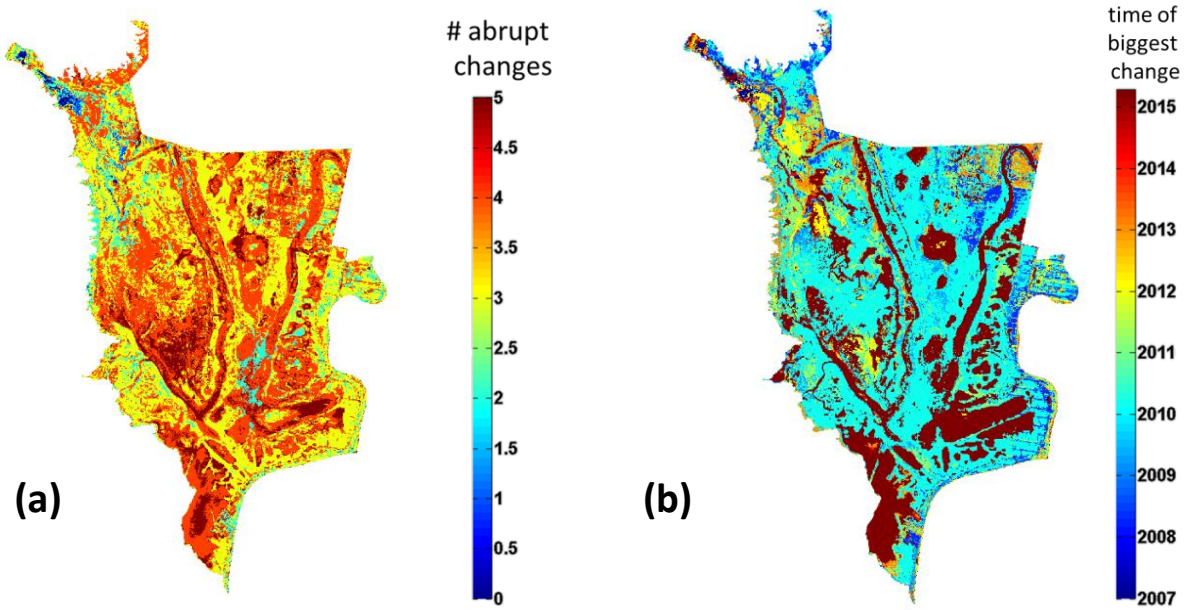

Figure 3. (a) Raster presenting the total number of abrupt changes detected from 01/01/2017 to 23/12/2016, (b) Raster presenting the time of biggest change from 2007 to 2016.

Figure 3 (a) shows that multiple changes are detected in the phenological cycles within the marshes during the last decade (2007 - 2016). The multiple changes in the trends of the vegetation cycle are probably related to changes in the hydroperiod regime in the Doñana marshes, which is mainly driven by the amount of precipitation.

The north-western part of the marsh is the part of the marshland that receives a more steady flow of water and thus it has a small number of abrupt changes (0-1 abrupt changes). The western edge of the marshes area called "Vera", which is the ecotone or transition zone between the sandy substrates and clay substrates of the marshes, is where ground water emerges providing a regular supply of water and grasslands are the dominant cover on it. The Vera appears to have 2 abrupt changes in this decade. Areas covered by the same vegetation community seem to have suffered a similar number of abrupt changes. Most of the pixels with value 3 changes correspond to areas called "vetas" and "paciles", i.e. subtly elevated areas as sandy-silt levies over the paleo-channels "caños". These areas, which are usually non-flooded for low and medium rainfall years, are occupied by *Arthrocnemum macrostachyum* and *Juncus subulatus vegetation communities.* The saltmarsh bulrush (*Bolboschoenus maritimus*) community, which lies in the transition zones between “paciles” and pond centers “Lucios”, seems to have suffered 4-5 abrupt changes, probably related to total annual rainfall and how rainfall levels influence the community productivity. Finally, the fact that paleo channels and depressions are shown to have the highest number of breaks, may be related to the fact that they are the most dynamic marshland areas showing big changes depending on whether they are flooded or not. For the analysed period, the total accumulated rainfall for 6 out of 11 flooding cycles has been under the average annual rainfall of 540mm (4 flooding cycles have annual rainfall in the 4th quartile and 2 in the 3d quartile; see Figure 4). This means that paleo channels and depressions areas have changed dramatically along the studied period.

Co-funded by the European Union

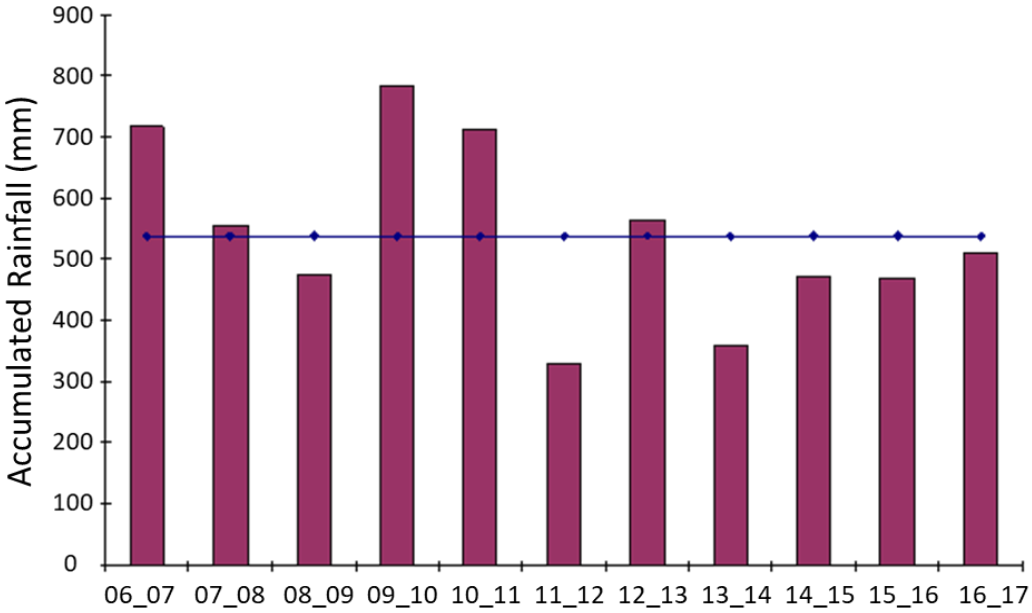

Figure 4. Accumulated Rainfall diagram for Doñana protected area from 2006 to 2017.

According to Figure 3(b), the biggest change for most of the marshland areas is detected between years 2010 and 2011. For the same period, there is a significant increase in the accumulated rainfall as it is evident in Figure 4. Thus, the time of the biggest change seems to correlate well with the recorded change of hydroperiod duration for these years. Apart from this, from Figure 3(b), it is possible to appreciate the difference in the "Caracoles" area state (north-eastern area), which was formerly an agricultural area and is gradually changing to a natural marsh. This state shows a different year for the greatest abrupt change. Also the north-western part with more permanent waters differs from the south-eastern part that is more seasonal.

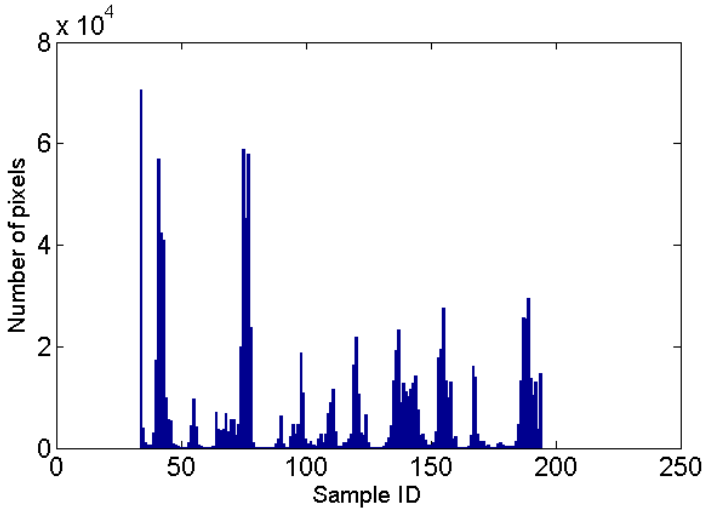

Figure 5. Histogram demonstrating the total number of pixels per sample ID, for which abrupt trend changes are detected.

Figure 5 indicates that most numerous abrupt trend changes are detected around sample IDs: 34, 41 and 75, which correspond to readings (2008, 179), (2008, 291), and (2010, 105) of the form (Year, Day from beginning of Year). On the other hand, Figure 4 indicates that most numerous biggest trend changes are detected around sample IDs: 75 and 189, which correspond to readings (2010, 105) and (2015, 104). Sample ID 75 is common in both histograms.

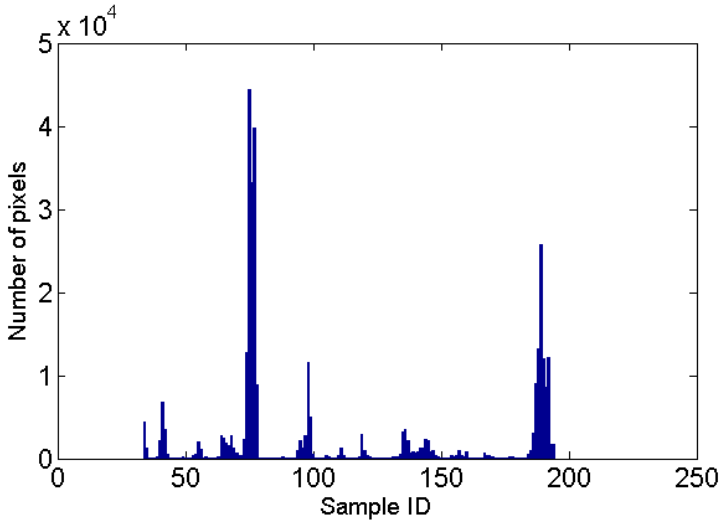

Figure 6. Histogram demonstrating the total number of pixels per sample ID, for which the biggest trend change is detected.

- Technological readiness level: 4
- Open source (Y/N): N, may be turned to open access, if resources allow so
- Commercial or proprietary (C/P): C
- Background or Foreground knowledge (B/F): F
- Partners: CERTH, CSIC

4.1.2 Time-series analysis of vegetation-cover response to residential development in dryland

Land-use (LU) transformations are known to affect their surroundings in terms of vegetation cover (VC). Dryland ecosystems are facing challenges of adaptation and reorganization in response to anthropogenic LU changes worldwide. Israel is one of the densest countries in the world, with the largest birth rate among the developed countries and increasing pressure for rapid residential development. The Negev, the southern arid part of Israel, is the largest land resource of the country, and government policy encourages redirecting growth to this region. Therefore, residential development is predicted to expand to this area, making it crucial to understand the effects of settlements development on their surrounding environment in general, and particularly on protected areas and natural reserves. We focused on assessing the impact of three types of residential settlements on VC through a multi-decadal time period in a dryland area. Three specific objectives were defined: (1) to form a time-series of 30 years (1987-2016) and produce a spatiotemporal trend map of VC change (VCC); (2) to analyze the main environmental factors that influence the VC trends; and (3) to determine to which extent and distance each settlement type in the region affects VC. Har-HaNegev is part the Negev desert and the elevation ranges between 290 and 860 m, with the southern part reaching higher elevations. Most of the VC is located in the stream network. Livestock grazing has been an ongoing phenomenon in this area for thousands of years and shaped the vegetation composition in this area. The selected site contains two towns, 11 single family homesteads, 10 livestock settlements and 10 government facilities, mostly military bases and a and a prison, as well as a protected natural reserves (Figure 7). Natural reserves occupy about 24% of the study area. The population concentration occupies about 0.3% of the area. Military training is allowed only in designated fire zones, which constitute about 60% of the study area. Anthropogenic activity is limited in the reserves and fire zones.

The study was conducted applying the following steps:

1. **Time series analysis:** we composed a series of Landsat images, all acquired during the summer. To analyze VCC, we selected the soil-adjusted vegetation index (SAVI) and applied it to every image. This vegetation index was designed to minimize soil brightness influence, which is very suitable for an arid environment with a low VC and a large soil background. The index was applied per-pixel such that $SAVI = \frac{(\rho NIR - \rho R)}{(\rho NIR + \rho R + L) * (1+L)}$; Where ρNIR and ρR

are the near infrared and red spectral bands of the sensor and L is an adjustment factor for reducing soil noise, and remains constant with a value of 0.5. A normalized SAVI time series was analyzed for assessing significant trends using the Contextual Mann-Kendall (CMK) significance test. This analysis assumes that conditions at a certain location will be similar to neighbouring locations, thereby allowing for spatial contextual information to apply resulting in Z score test statistic for every pixel. A pixel with a positive Z score indicates a positive trend whereas negative scores suggest a negative trend. We considered CMK's *p-values*≤0.05 to be significant (Z score <-1.96 and >1.96). The resulting map provided a Z score for each pixel, indicating the magnitude and direction of the VC trend through time.

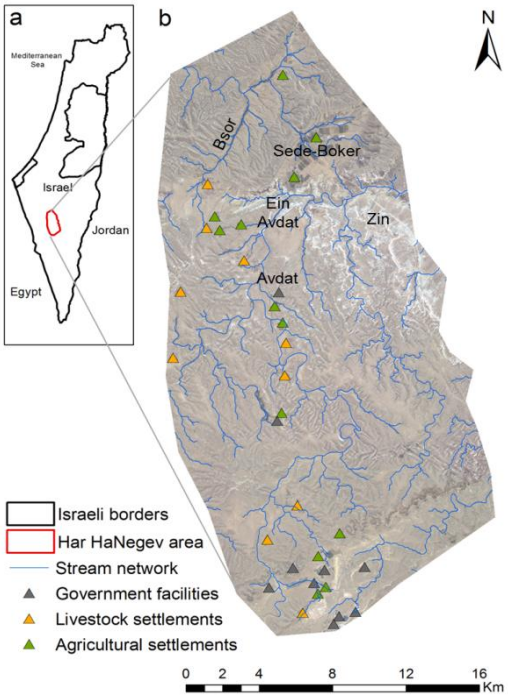

Figure 7: Har HaNegev study site

2. **Spatial cluster analysis:** We examined spatial trends as the degree to which the CMK Z score value of a specific pixel is similar to that of a neighbouring pixel. Spatial phenomena exhibit non-stationarity and spatial dependence among features and often cluster in space. We applied the spatial Getis-Ord G_i^* statistic (hot-spot analysis) to analyze the location-related tendency in points attributes and evaluate the similarity (high or low) between each point and its surrounding features. The clusters were then spatially correlated with distance from streams and elevation, using geographically weighted regression (GWR).

3. **Environmental variables effect on vegetation spatial distribution:** We analyzed selected variables that may influence VCC spatially, including elevation, slope, aspect, distance from streams, fire zones, natural reserves, temperature, soils, lithology, land cover, populated areas, and distance from populated areas.

4. **Trends in VC as a function of distance from residential concentrations:** We divided the residential concentrations in the study area into three settlement types and examined their effect on VC in various distances. Each residential concentration was represented by a centroid. Then, buffers were generated for the following distances from centroids: 100, 150, 200, 250, 500, 1000, 1500, 2000, 2500, 3000, 3500, 4000, and 4500 meters.

We used 12,788 significant Z score points across the study site. For each distance from the centroids of each settlement type, the Z score points were extracted and the ratio between positive and negative values was computed. These ratios were compared between the settlement types. The analysis of variance (ANOVA) was calculated to study the differences between the various distances for each settlement type and the differences between the various settlement types for each distance.

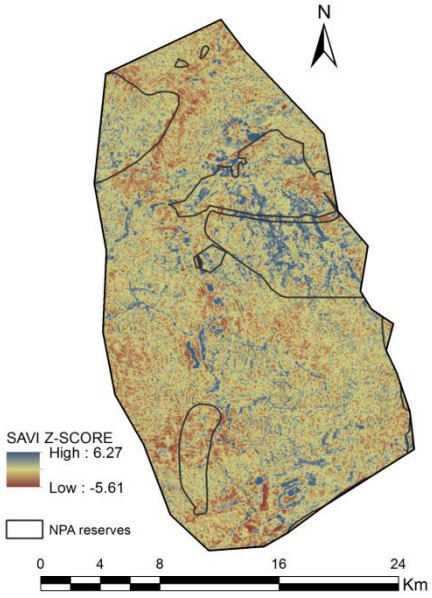

Figure 8: A temporal Z score map of vegetation cover trends

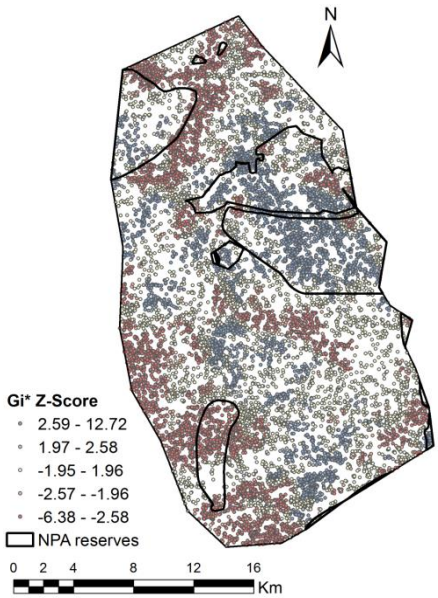

Figure 9: A spatial Z score map of clustering vegetation-cover change.

Figure 8 shows the results of the CMK time series analysis, representing the temporal trend of each pixel. Three major patterns were spatially explicit: 1) the areas of the NPA natural reserves (protected areas) had the highest concentration of Z score values, representing positive trends of VCC through 30 years of analysis; 2) areas that have experienced grazing during this period show very low, significantly negative Z score values, indicating a sharp decrease in VC and a degradation pattern; 3) agricultural activities provide conditions for increase in VC. Figure 9 provides an additional spatial cluster analysis. The spatial extents of the original Z scores show high positive clustering in the Zin NPA reserve (a protected area), and high negative clustering in areas prone to livestock activities. The clustering patterns (both positive and negative) are mostly confined to the stream network ($R_{adjusted}$=0.93). The spatial clustering was also responsive to the elevation gradient ($R_{adjusted}$=0.98). The results of the GWR analysis show that the most influencing explanatory variable on VCC temporal trend, was soil cover ($R_{adjusted}$=0.62), followed by elevation and LST ($R_{adjusted}$=0.59, 0.52, respectively).

The negative-positive Z score ratio values that were calculated for each distance from the settlements' centroids were plotted in Figure 10. The ratio for the entire study area was composed of 4502 positive sample points and 8286 negative points, situating the overall ratio at -0.84. This ratio reflects the general negative VCC trend in the Har HaNegev area. As an additional reference, we calculated the ratio for the NPA natural reserves, where human disturbances are kept to a minimum and found a ratio of 0.17. The ratio for the combined settlements started positively at low distances from the centroids, gradually declining with distance; after about 250 m the ratio

Co-funded by the
European Union

dropped below 0, indicating a negative ratio in VCC. After 1000 m the ratio reached the point of overall ratio and kept declining until reaching stability at 1500 m. For agricultural settlements, the trend was similar, however the ratio values were higher than for the average. At around 3000 m, the ratio merges with the overall area's ratio. For livestock settlements, the ratio started with >0 values in short distances from the centroids, and dropped below 0 at 500 m. The ratio values at 1000 m were significantly lower than both the average and all settlements ratios, and stayed low for up to 4500 m, moderately increasing at 2000 m and as distance grew longer. For government facilities, the ratio was always <0, merging with the overall ratio at a distance of 1000 m.

Co-funded by the European Union

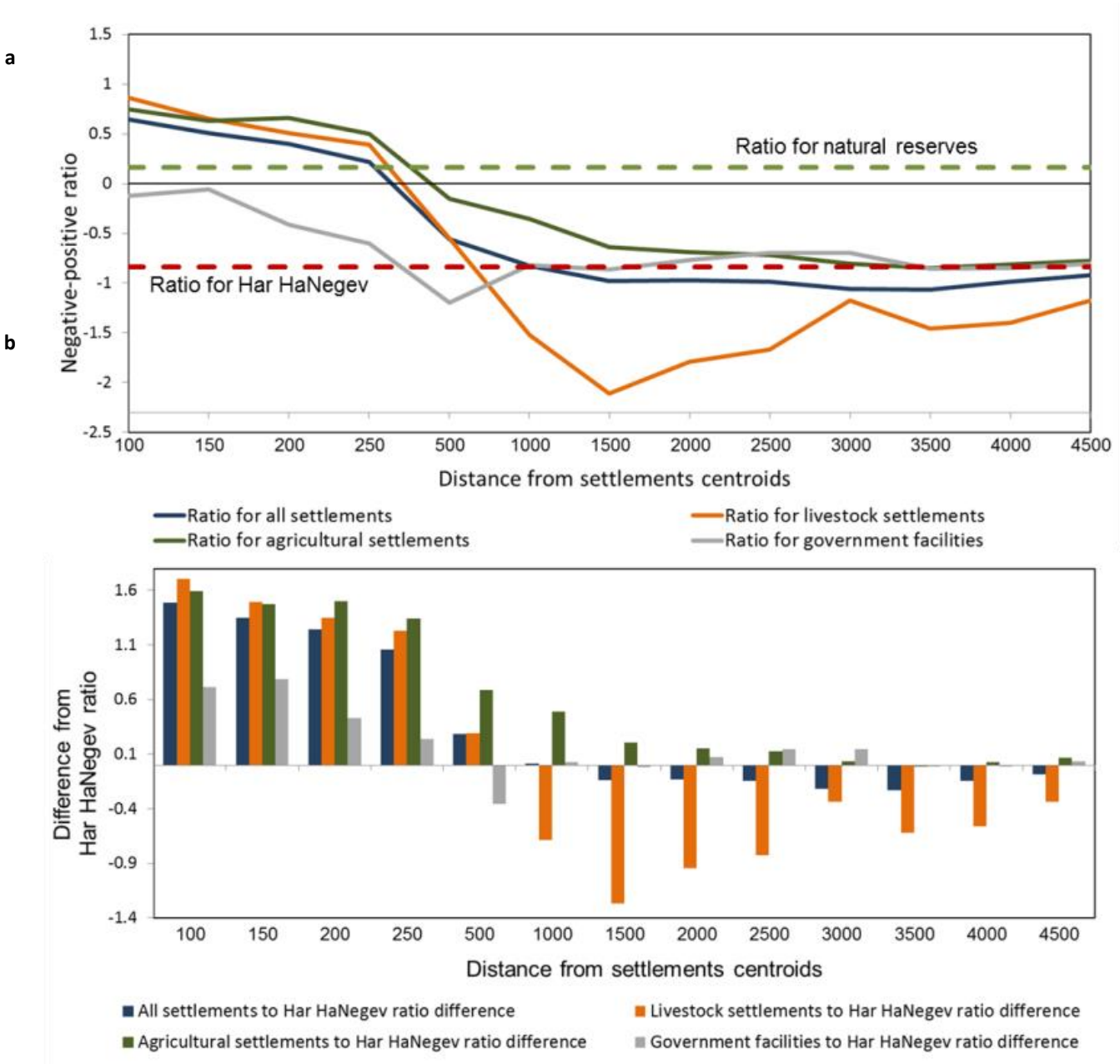

Figure 10: A negative-positive temporal Z score ratio as a function of distance from the settlements' centroids. a) Ratios relative to 0; b) Ratios as differences from the overall Har HaNegev ratio.

The residential concentrations types vary greatly in their influence on VCC in Har HaNegev. As the results show (Figure 10, Tables 2, 3), the overall negative-positive Z score ratio is negative, meaning that in average, the Har HaNegev area is experiencing a decrease in VCC over the last 30 years. However, agricultural activities near centroids increase VC, due to intensive cultivation within and close to the settlements. The ratio gradually decreases and turns negative with distance. The livestock settlements have positive impact on VCC near the settlements' centroids that gradually decrease with distance. They show very low Z score ratios at increasing distances, with a low climax at 1500 m from the centroids and their effect still exists at a distance of 4500 m. The areas of government facilities have little portions of VC; therefore, their ratio is always negative and significantly different from that of agricultural and grazing settlements. Our findings indicate that in NPA natural reserves, the total ratio was positive.

Distance from centroid	Agricultural settlements	Livestock settlements	Government facilities
100	a	a	b
200	a	a	b
500	a	b	c
1000	a	b	c
1500	a	b	c
3000	a	b	a
4500	a	b	a

Table 3: Significant differences between Z scores for settlement types for each selected distance from settlements' centroids.

Settlement type	100	200	500	1000	1500	3000	4500
Agricultural settlements	a	a	b	bc	d	d	d
Livestock settlements	a	a	b	cd	d	d	c
Government facilities	a	a	a	a	a	a	a

Table 4: Significant differences between the Z scores for the selected distances of each settlement type.

Protected natural reserves are known to be an effective conservation instrument to prevent natural vegetation loss compared to their adjacent environment. Successful preservation of the natural landscape and mitigation of vegetation clearing, relies on reducing human pressure and applying suitable management strategies. Negative VCC in drylands may have irreversible effects that influence the vegetation pattern and organization in space, thus altering ecosystem structure and pattern by changing the landscape attributes and interrelationships between vegetation and various abiotic factors (climate, soil characteristics, etc.). After decades of vegetation clearing in Har HaNegev, considerable management strategies and measures should be implemented in order to begin a conservation and restoration process and improve the environmental status of this site. Governmental and municipal regulations are needed before the areas adjacent to natural reserves become more populated and correspondingly, degradation processes will further expand.

4.1.3 Cloudiness time series analysis on La Palma Island (Canary Islands)

The whole island of La Palma represents a UNESCO worldwide biosphere reserve since 2011. The central part of the island is preserved as a national park "Caldera de Taburiente" established as early as 1954.

La Palma is the north-western island of the Canary Islands, positioned in the Atlantic Ocean off the coast of north-western Africa: it is a relatively young volcanic island. Climatically, La Palma exhibits moist winter and dry summer but, the climatic conditions differ considerably within the island with strong differences in the annual temperatures strongly influenced by the trade winds blowing from the Northeast. This creates a climatic divide of the island. The windward side is generally more humid receiving the most precipitation and higher cloud cover, while the leeward side is dryer and has more insolation. However, above the trade wind-induced cloud zone a thermal inversion layer exists, generally exhibiting dry conditions and possessing the possibility of snow and ice in winter (García-Herrera et al., 2001).

Cloudiness is one of the ECVs regulating Earth radiation budget and water cycle and exerts a strong feedback on climate change. The variable properties of clouds determine their profound effects on radiation and precipitation. They are influenced by and in turn influence the motion of the atmosphere on many scales. They are affected by the presence of aerosols, and modify atmospheric composition in several ways, including the depletion of ozone when they form in the polar stratosphere. The feedback from changes in clouds remains one of the most uncertain aspects of future climate projections and is primarily responsible for the wide range of estimates of climate sensitivity from models. According to the Global Climate Observing System (GCOS) which has developed Climate

Monitoring Principles that set out general guidelines to achieve observations and addressing key satellite-specific operational issues, observations of cloud properties are needed for improved understanding and quantification of both local and larger-scale cloud-related processes, for climate monitoring, for validation and development of numerical models and for their emerging use with these models in data assimilation (GCOS Implementation Plan, 2016)

Cloudiness is a major ecosystem driver of La Palma supporting ecosystem services such as water provision.

The work focuses on assessing the effectiveness of free available Landsat and Sentinel-2 images time series to obtain cloudiness time series and produce on La Palma island a map of frequency for cloud presence.

Three different areas, 1 km x 1 km size, distributed on the North-West (AREA_1_NW), North-East (AREA_2_NE) and South (AREA_3_S) of the whole island (Figure 12), were investigated and the cloudiness against time was plotted.

This represents a preliminary study so short time series of satellite images were considered.

Images from two different sensors, Landsat 8 OLI/TIRS and Sentinel-2A, were considered: multi-source data could be used for improving the temporal resolution of time series.

For the period from August 2015 to December 2015, a series of Landsat 8 OLI/TIRS images, two images per month, with a total amount of 10 images was composed (see Table 5). Differently from the usual request, this time even satellite images with cloud presence were considered (Figure 11). The images, at 30 meters spatial resolution, were projected in WGS84/UTM28N.

Year	Date in DD/MM format									
2015	09/08	25/0	10/0	26/0	12/1	28/1	13/1	29/11	15/12	31/12

Table 5. Dates of Landsat 8 OLI/TIRS images considered.

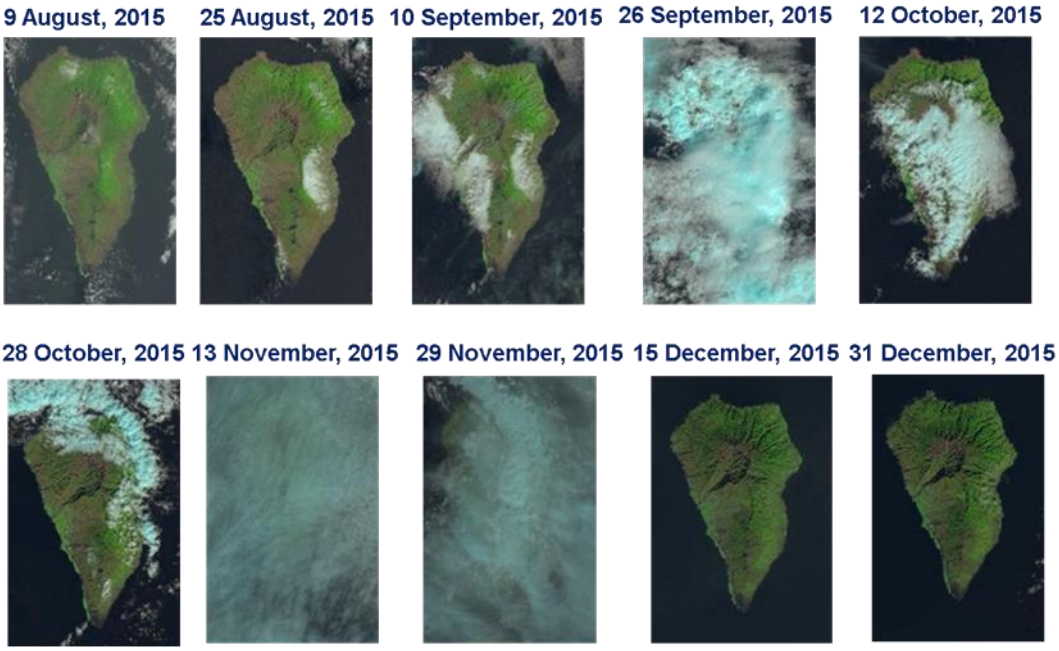

Figure 11. Landsat 8 OLI/TIRS time series. Real Color Composition.

To extract the cloud masks for each Landsat 8 OLI/TIRS image, first, a radiometric calibration in reflectance at Top Of Atmosphere (TOA) was applied to multispectral bands in the visible, near infrared, short wave infrared and cirrus whereas a radiometric calibration in brightness emissivity was requested for thermal bands. Then the "Fmask"

algorithm (Zhu et al., 2015) was applied for each pixel producing a cloud mask (for additional information see deliverable D4.3, paragraph 3.3).

Cloudiness values in the time series, defined as the percentage of pixels corresponding to cloud for a given location, were plotted for the three areas distributed on the scene (Figure 12).

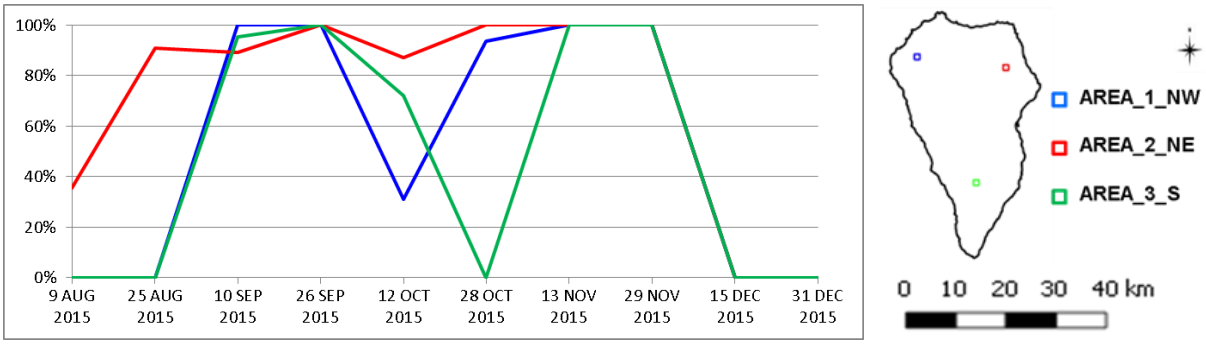

Figure 12. Cloudiness values in the Landsat 8 OLI/TIRS time series, in three different areas.

According to the well-known dominant winds from Northeast direction, cloudiness values for "AREA_2_NE" results at the highest values for a longer time respect to the other two areas.

Finally, combining all the 10 cloud masks a map of the frequency of cloud presence for each pixel was obtained: the higher values correspond to the higher frequency presence (Figure 13).

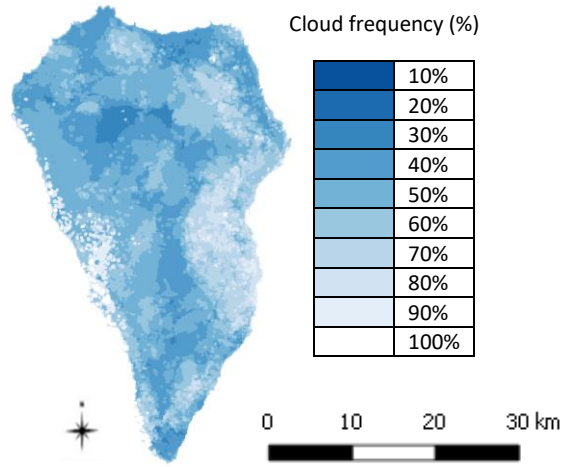

Figure 13. Cloud presence frequency map from Landsat 8 OLI/TIRS time series.

For the period from December 2015 to April 2016, a series of cloud masks from Sentinel-2A images, one image per month, with a total amount of 5 images was composed (see Table 6). The cloud masks associated to Sentinel-2 images were free available and downloadable from Google Earth Engine cloud platform as monthly cloud masks product and they result from the OR among different temporal acquisitions in a month (Figure 14). The masks, at 60 meters spatial resolution, were projected in WGS84/UTM28N.

Year	Date

2015	December				
2016		January	February	March	April

Table 6. Dates of Sentinel-2 A images considered.

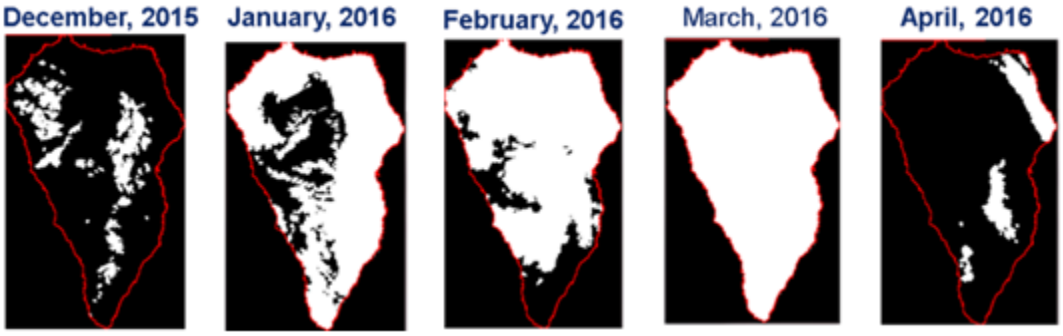

Figure 14. Monthly cloud masks from Sentinel-2A time series.

The cloud masks of Sentinel-2 products have a coarse spatial and temporal resolution with respect to the Landsat cloud masks, even though it requires no any other processing step and it is ready for the analysis.

The cloudiness values in the time series were plotted for the three areas distributed on the scene (Figure 15).

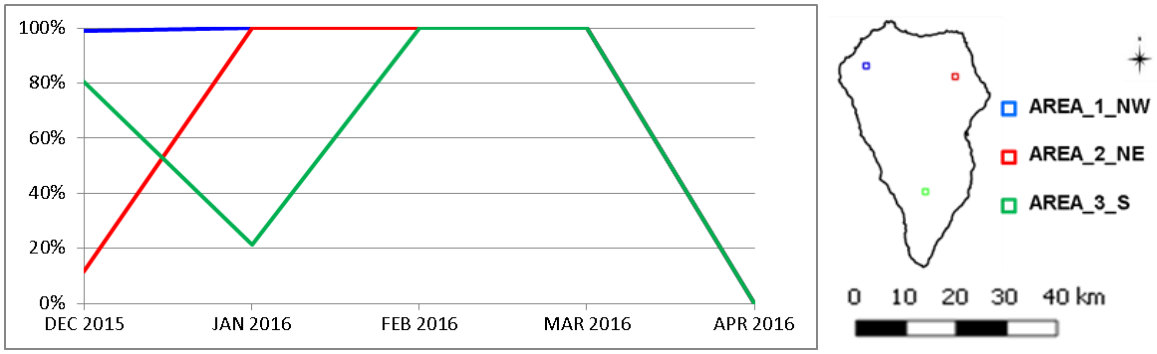

Figure 15. Cloudiness values in the Sentinel-2A time series, in three different areas.

In this period the North of the island results having high cloudiness values for a longer time.

Finally, combining all the Sentinel- 2 derived cloud masks, a map of the frequency of cloud presence for each pixel was obtained: the higher values correspond to the higher frequency presence (Figure 16).

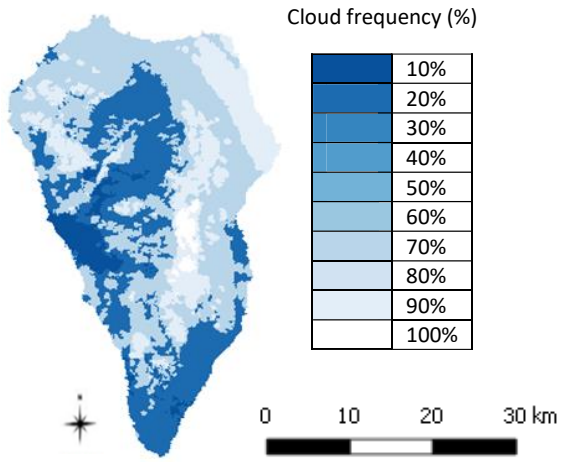

Figure 16. Cloud presence frequency map from Sentinel-2A time series.

4.2 Coastal and marine

4.2.1 Status and evolution of food provision service: sea surface temperature time series analysis using STL and EOF on Mediterranean LME

The Mediterranean Sea is a Large Marine Ecosystem and the largest enclosed sea in the world. Its 46,000 km coastline spans many countries across three continents (Europe, Africa and Asia). Rich in marine life and other features, it is estimated to host between 4 per cent and 18 per cent of the planet's macroscopic marine organisms. It provides many valuable goods and ecosystem services to society, including coastal protection, food, tourism, recreational opportunities and many more. The Mediterranean Sea's fish and other living resources supply the fishery and aquaculture sectors, which may be threatened by human activities, especially under changing climate conditions. Monitoring seawater conditions is therefore of paramount importance for scientists and the Marine Protected Area managers, who are involved in the conservation of marine natural resources. For example, changes in seawater temperature affect the delivery of ecosystems services such as food provision, as seawater temperature influences the activity and health of fish, including their feeding, reproduction, movement and distribution.
Through the use of Earth Observation (by examining over 12,000 images) and other tools, the ECOPOTENTIAL project team has observed that in the near future, the Mediterranean Sea will become progressively warmer, affecting the movement and distribution of fish and thereby generating new food-provision scenarios.

The main factors controlling the fish growth are related to fish vitality, fish weight, and aquaculture technical constraints, and more in general to the local physical-chemical and biological properties of the water that affect directly the fishes or their habitat. Physical, chemical and biological characteristics of seawaters are therefore the primary descriptors to assess the spatial and temporal dimensions of marine ecopotential productivity performances in terms of fish growth. Among these seawater characteristics, seawater temperature is the Essential Variable (EV) to characterize fish vitality and thus the best candidate to marine food provisioning service assessment, as it influences many other parameters and as consequence it reflect an important part of the entire life cycle of marine organisms.
The knowledge of spatial and temporal dimensions of marine ecopotential productivity performances in terms of fish growth can be evaluated also by multitemporal analysis of sea surface temperature data.

Here the Seasonal Trend Decomposition (STL) analytical technique divides up a time series for the period 1982-2016 collected from multi-sensor satellite data available in Copernicus Marine Environment Monitoring Service (CMEMS) into three components, namely trend, seasonality and remainder and is able to shows intra-annual

seasonal variability. In the Mediterranean case this technique is applied on time series of daily SST estimated.

The signal decomposition analysis (see paragraph 3.2.2) was done using R software language with the support of packages 'rtsa' (see paragraph 3.5), 'stlplus' (Hafen, 2016) and 'Kendall' (McLeod, 2011).

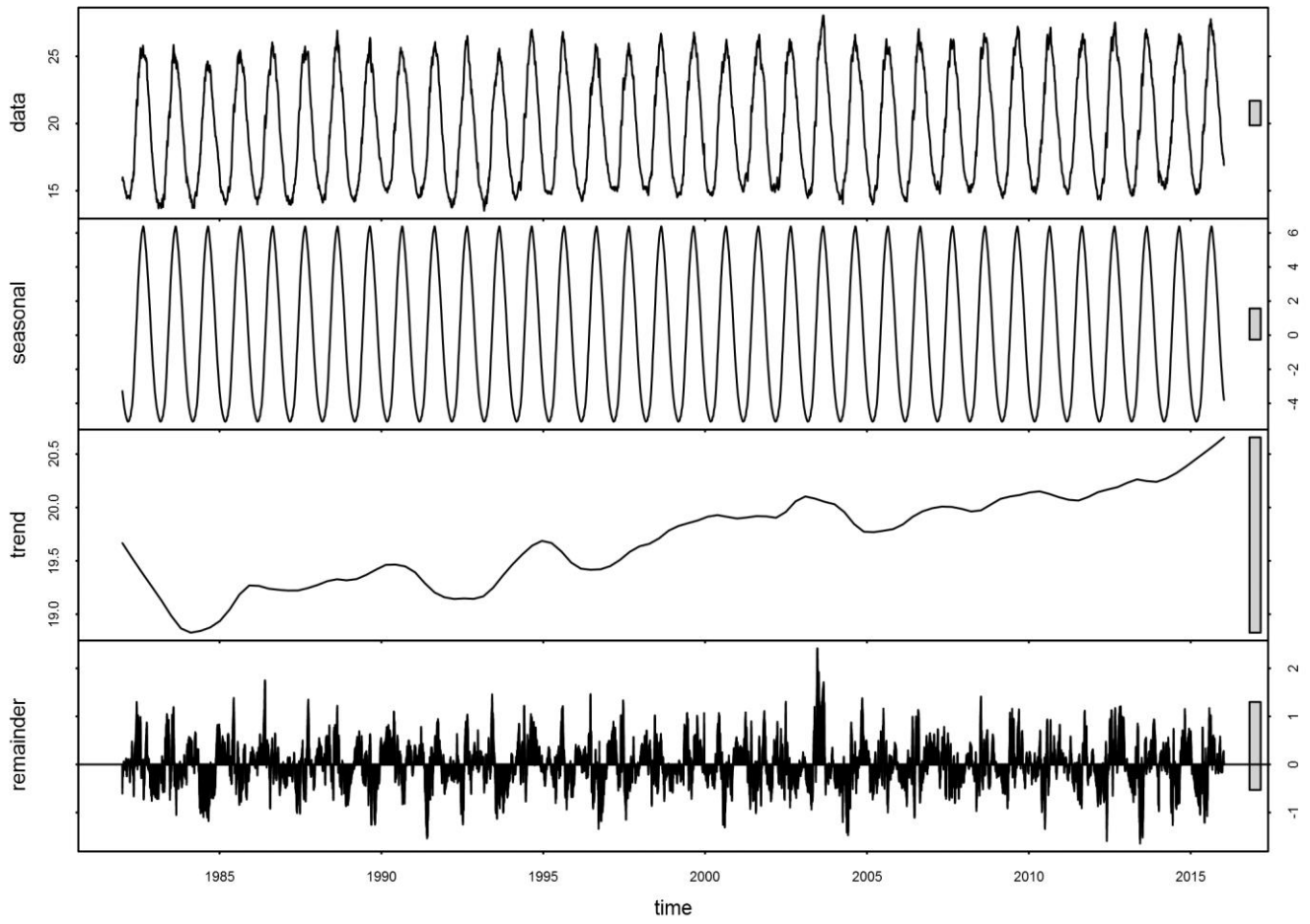

Figure 17. Result from the signal decomposition analysis computed from the Mediterranean SST temporal profile of a single pixel. Original data, seasonal component, trend component and remainder are displayed. Units are in Celsius degrees.

Figure 17 shows the seasonal component and trend component obtained applying STL signal decomposition analysis to the SST temporal profile of a single pixel, representing the average conditions of the entire Mediterranean basin. Seasonal amplitude was found to be about 11 °C, while trend component show a constant positive slope of about 0.03 degrees/year. By computing STL for each pixel it is possible to obtain a spatial representation of the information derived from the signal decomposition analysis, like the seasonal amplitude (Figure 18) and the trend slope (Figure 19).

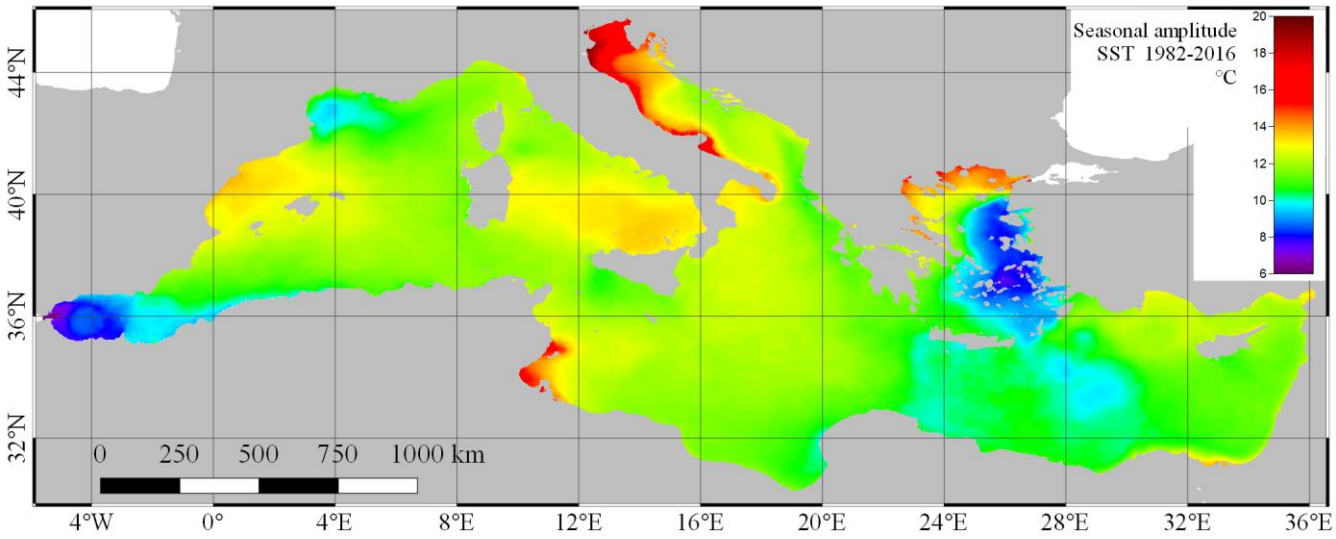

Figure 18. STL analysis: SST seasonal amplitude for the period 1982-2016.

STL analysis revealed a higher intra-annual seasonal variability in the Adriatic Sea, Thracian Sea and the Gulf of Gabes, primarily due to the shallow waters characteristics. On the contrary, lower intra-annual variability was found in the easternmost part of the Aegean Sea and in the Strait of Gibraltar, the latter related to the water exchange with the Atlantic Ocean (Figure 18).

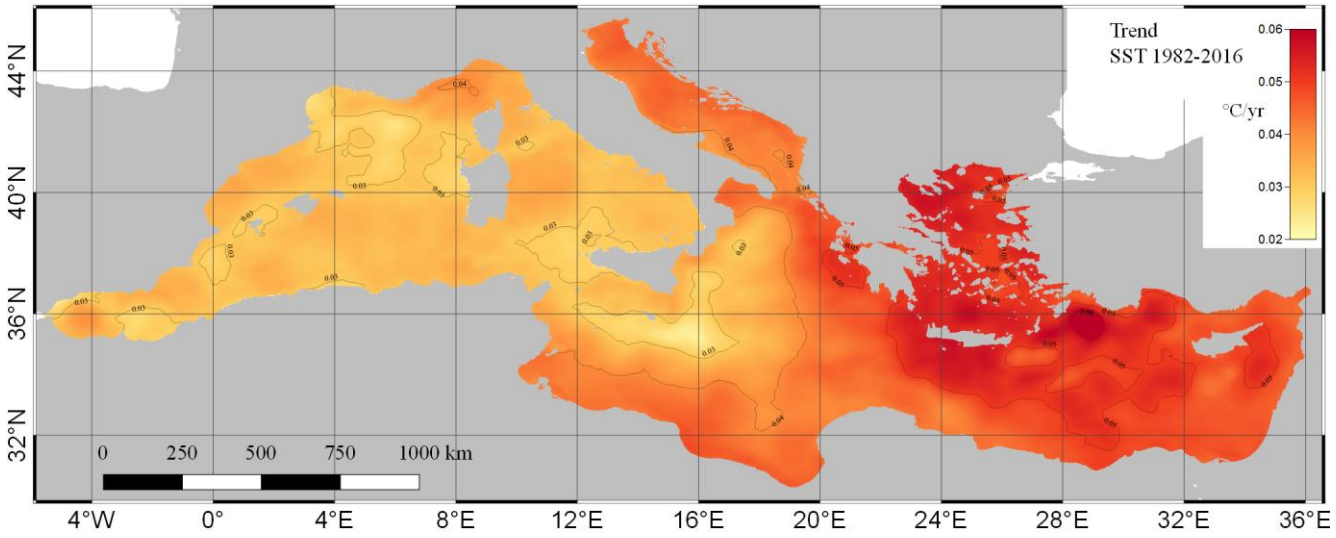

Figure 19. STL analysis: SST trend using Seasonal Trend Decomposition analysis for the period 1982-2016 (Mann - Kendall test resulted in p-values<= 0.0001 for all the pixels)

Resulting trend signal from STL analysis (Figure 19) revealed an increasing SST trend in the all Mediterranean Sea basin for the period 1982-2016 ranging among 0.02 and 0.08 degrees/year, corresponding to an average value of 1.4 Celsius degrees for the entire considered period. The higher increasing trend is located in the eastern part of the basin, mostly in the Aegean Sea. The multitemporal analysis of SST represents an assessment of and ecosystem conditions (thermal condition in the form of an EV) which is fundamental to current food provision service evaluation (see task 4.3.4), and provides the basis for assessing its evolution: the results from trend analysis can be combined with a fish growth model to generate future scenarios of fish growth rates. Additional results from multitemporal analysis of SST time series will be described in D8.2, which is expected to show past and current change of ecosystems state and ecosystem services delivered.

EOF is a principal component technique that shows the spatial patterns of variability of considered data, their time variation and the importance of each pattern on the basis of variance (see paragraph 3.3.1). Each principal component loading pattern is named mode, and can be both represented in its spatial dimension and temporal dimension (expansion coefficient). The EOF analysis was done using R software language with the support of packages 'rtsa' (see paragraph 3.5) and 'sinkr' (Taylor, 2017).

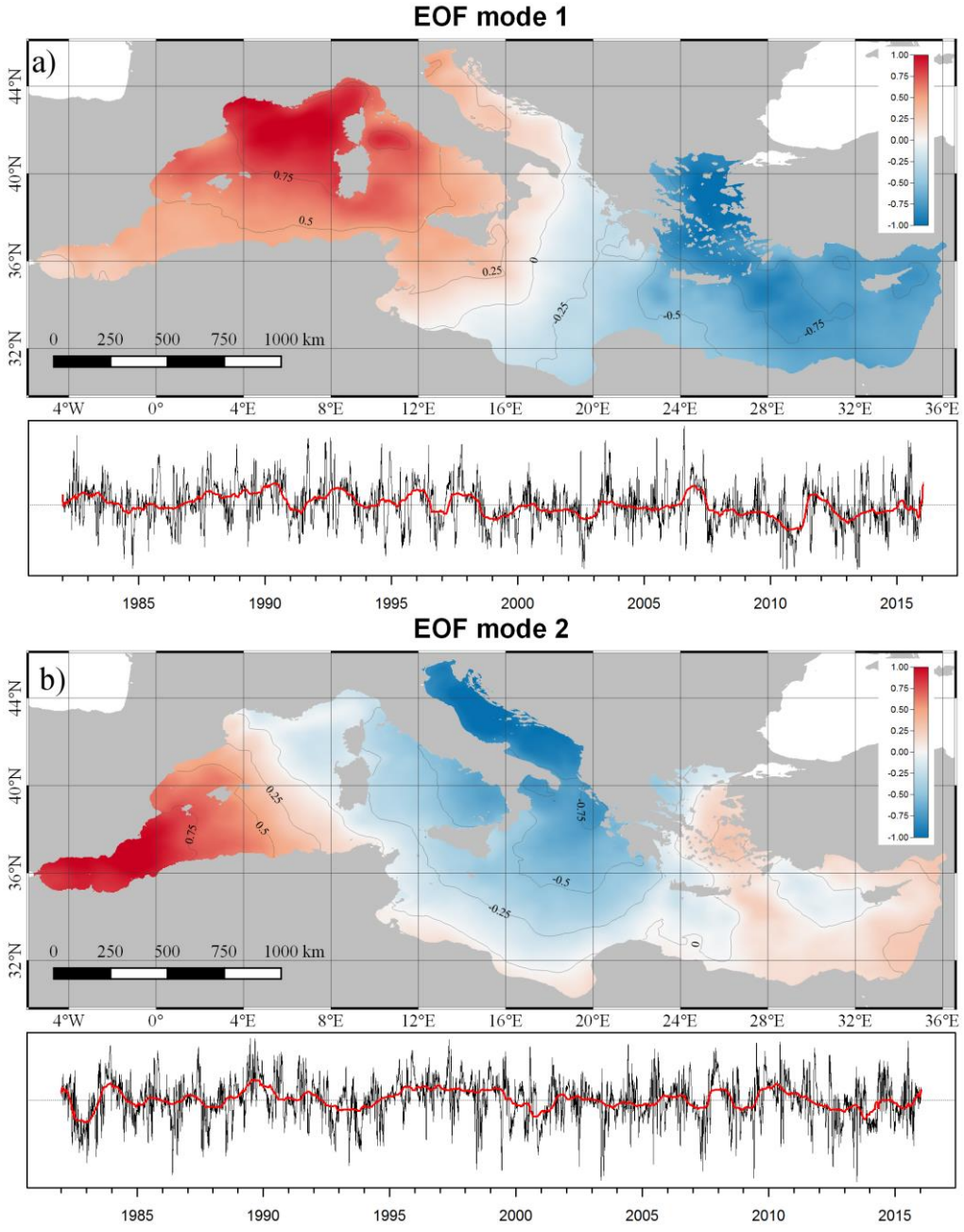

Figure 20. Modes and Expansion Coefficients resulting from EOF analysis: (a) EOF mode 1 and the correspondent EC; (b) EOF mode 2 and the correspondent EC

The first 4 EOFs generated from deseasonalized SST time series contain over 80% of the overall data variability (in Figure 20 are showed only the first two modes).

EOF results depend on the data used as input (e.g. EOF mode 1 generated from anomaly SST or from deseasonalized SST result in different spatial patterns) and this technique does not have the ability to split the contribution of different forcings into the different modes (modes are mixed contributions) and therefore modes interpretation could be difficult. In the Mediterranean case spatial patterns are strongly related to atmospheric circulation, but a more expert based knowledge is required for a more complete interpretation of the patterns.

SOM multitemporal analysis 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 of input. It is a very complex analysis whose results strongly require to be interpreted in the light of other variables combining with the input variable.

In our case we decided to produce 4 SOMs that can represent 4 of the main behaviours due to main forcing factors (e.g. fundamental circulation pattern and the related longitudinal gradient).

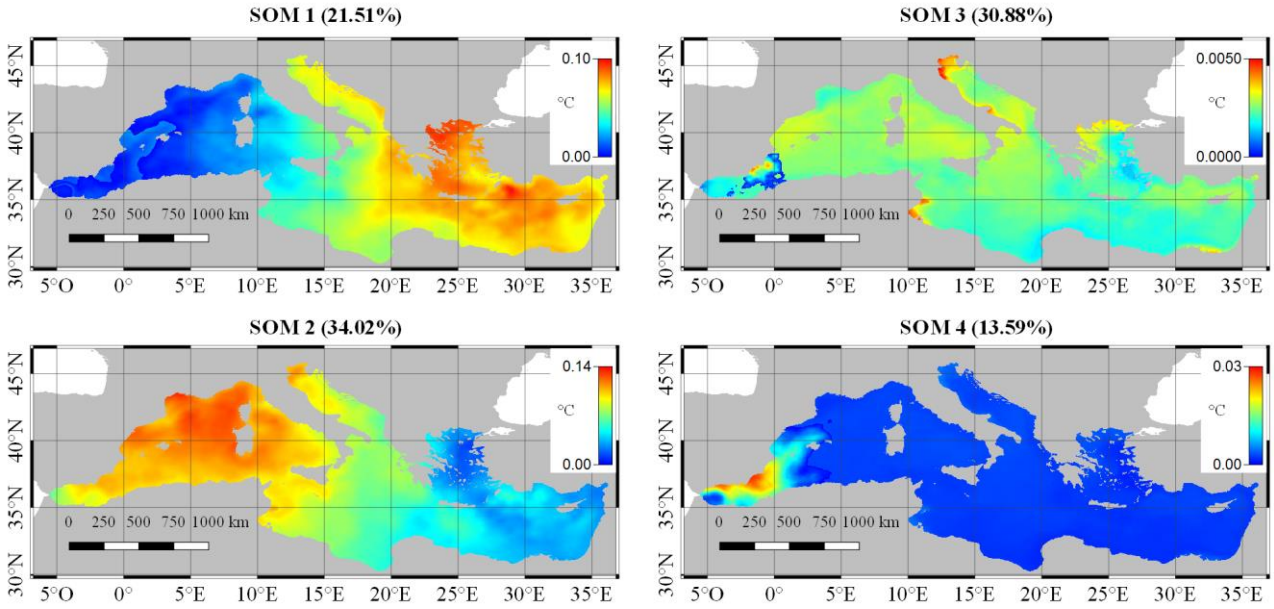

Figure 21. Spatial representation of the four Mediterranean Self Organized Maps (deseasonalized)

5. Results

The aim of the report is twofold. On one hand (technical perspective) providing an overview of the available or developed techniques to perform multitemporal analysis on continuous data (e.g. environmental variables, indicators and indexes showed in D4.2) that could feed the EODESM and the project exploitation platforms, and supporting activities of ecosystem modelling, ecosystem service assessment. Among the results, a suite of techniques described in this document is provided as a software package, for which the TRL is supplied.

As some of the project partners may not be in a position to provide their modules or to develop software packages within the ECOPOTENTIAL project, not all the described techniques are made available.

On the other hand (knowledge perspective) the reports presents some applications of the described available techniques in Ecopotential sites in order to show which techniques should be used according to the phenomenon you want to observe. The result of this second objective is a table which synthesizes for selected variables which techniques and algorithms could be used in order to obtain the relevant information.

Moreover the report represents a step towards the exploitation of the large amount of available multisource temporal series of EO data (amount that is expected to keep on growing very fast due to the Sentinels constellation)

in the perspective of ecosystem services assessment.

5.1 Comparison of analytical techniques: strengths and weaknesses

An overall comparison on the strengths and weaknesses for some of the different multitemporal analysis methodologies, coming from research studies (Filipponi et al., 2017), is reported in Table 7. There are still few comparisons over analytical methods for multitemporal analysis as well as few requests of data analytics to extract information from large EO products datasets.

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 (almost) 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
X-11 and X-13ARIMA-SEATS	• Capable of extracting the seasonal, trend and residual component • Capable of dealing with varying seasonal signal throughout the years • 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 • Only supports monthly time series • Compute-intensive • Computation allowed for a limited number of observations in time • Work on single pixel temporal profile, not accounting for spatial variability • Spatial variability may result in uneven patterns, probably due to different model setup for different pixels • User need to define the time windows for the extraction of 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 7. Strengths and weaknesses for some of the described analytical methodologies.

6. References

1. Alvera-Azcárate, A., Barth, A., Sirjacobs, D., & Beckers, J.-M. (2009). Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF. Ocean Sci., 5, 475-485.
2. Baglama J., & Reichel, L. (2005). Augmented Implicitly Restarted Lanczos Bidiagonalization Methods. SIAM J. Sci. Comput. 27(1), 19-42.
3. Beckers, J.-M., Barth, A., & Alvera-Azcarate, A (2006). DINEOF reconstruction of clouded images including error maps - application to the Sea-Surface Temperature around Corsican Island. Ocean Science, European

Geosciences Union, 2 (2), pp.183-199.

4. Beckers, J.-M., Barth, A., Tomazic, I., & Alvera-Azcárate, A. (2014). A method to generate fully multi-scale optimal interpolation by combining efficient single process analyses, illustrated by a DINEOF analysis spiced with a local optimal interpolation. Ocean Sci., 10, 845-862.
5. Bjoernsson, H. & Venegas, S.A. (1997). A manual for EOF and SVD analyses of climate data, McGill University, CCGCR Report No. 97-1, 52pp.
6. Chen, J., Jonsson, P., Tamura, M., Gu, Z., Matsushita, B., & Eklundh, L. (2004). A simple method for reconstructing a high quality NDVI time series data set based on the Savitzky-Golay filter, Remote Sens. Environ., 91, 332-344.
7. Cleveland, R. B., Cleveland, W. S., & Terpenning, I. (1990). STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics*, *6*(1), 3.
8. Cleveland, W.S., & Grosse E. (1990). Fitting Curves and Surfaces to Data. Monterey, CA: Wadsworth Advanced Books and Software.
9. Dagum, E. B. (1978). Modelling, forecasting and seasonally adjusting economic time series with the X-11 ARIMA method. Journal of the Royal Statistical Society. Series D (The Statistician), 27(3/4), 203-216.
10. De Oliveira, J.C., & Epiphanio, J.C.N. (2012). Noise reduction in MODIS NDVI time series data based on spatial-temporal analysis. In Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International, pp. 2372-2375.
11. De Oliveira, J.C., Epiphanio, J.C.N., & Renno, C.D. (2014). Window regression: A spatial-temporal analysis to estimate pixels classified as low-quality in MODIS NDVI time series. Remote Sensing, 6(4), 3123.
12. DeVries, B., Pratihast, A.K., Verbesselt, J., Kooistra, L., & Herold, M. (2016). Characterizing Forest Change Using Community-Based Monitoring Data and Landsat Time Series. Vadrevu KP, ed. PLoS ONE. 11(3), e0147121.
13. Diaz-Delgado, R. (2010). An Integrated Monitoring Programme for Doñana Natural Space: The set-up and Implementation. In: Conservation Monitoring in Freshwater Habitats: A Practical Guide and Case Studies, pp. 375-386.
14. Eastman, J. R. (2016). TerrSet Manual. Worcester, Massachussetts: Clark Labs, Clark University.
15. Falcieri, F.M., Benetazzo, A., Sclavo, M., Russo, A., & Carniel, S. (2014). Po River plume pattern variability investigated from model data, Continental Shelf Research, vol. 87, pp. 84-95.
16. Filipponi, F., Valentini, E., & Taramelli, A. (2017). Sea Surface Temperature changes analysis, an Essential Climate Variable for Ecosystem Services provisioning. 2017 9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Bruges (Belgium) 27-29 June 2017, IEEE Conference Publications.
17. Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C., & Chen, B.C. (1998). New Capabilities and Methods of the X-12-ARIMA Seasonal Adjustment Program (with discussion). Journal of Business and Economic Statistics, 12.
18. Forsythe, G.E., Malcolm, M.A., & Moler, C.B. (1977). Computer Methods for Mathematical Computations. Wiley.
19. García-Herrera, R., Gallego, D., Hernandez, E., Gimeno, L., & Ribera, P. (2001). Influence of the North Atlantic Oscillation on the Canary Island precipitation. J. Clim., 14: 3889-3903.
20. García Novo, F., & Marín Cabrera, C. (2005). Doñana: Agua y Biosfera. Ministerio de Medio Ambiente, Sevilla, 354 pp (in Spanish).

21. GCOS 2016 Implementation Plan (2016). The Global Observing System for Climate: Implementation Needs, http://gcos.wmo.int.
22. Gerber, F., Furrer, R., Schaepman-Strub, G., de Jong, R., & Schaepman, M.E. (2016). Predicting missing values in spatio-temporal satellite data. arXiv preprint arXiv:1605.01038.
23. Gilbert, R.O. (1987) . Statistical Methods for Environmental Pollution Monitoring. Wiley, NY.
24. Golyandina, N., & Osipov, E. (2007). The "Caterpillar"-SSA method for analysis of time series with missing values, J. Stat. Plan. Infer., 137, 2642-2653.
25. Gómez, V., & Maravall, A. (1996). Programs TRAMO (Time series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series) Instructions for the User. Working Paper, Banco de España.
26. Hafen, R. P. (2010). Local regression models: Advancements, applications, and new methods. West Lafayette, Indiana: Purdue University, PhD dissertation, pp. 279.
27. Hafen, R. (2016). stlplus: Enhanced Seasonal Decomposition of Time Series by Loess. R package version 0.5.1. https://CRAN.R-project.org/package=stlplus.
28. Hopkins, J., Lucas, M., Dufau, C., Sutton, M., Stum, J., Lauret, O., & Channelliere, C. (2013). Detection and variability of the Congo River plume from satellite derived sea surface temperature, salinity, ocean colour and sea level. Remote sensing of environment, 139, 365-385.
29. Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.C., Tung, C.C., & Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. In *Proceedings of the Royal Society of London A: mathematical, physical and engineering sciences* (Vol. 454, No. 1971, pp. 903-995). The Royal Society.
30. Jamali, S., Jönsson, P., Eklundh, L., Ardö, J., & Seaquist, J. (2015). Detecting changes in vegetation trends using time series segmentation. Remote Sensing of Environment, 156, 182-195.
31. Jonsson, P., & Eklundh, L. (2002). Seasonality extraction by function fitting to time series of satellite sensor data. IEEE T. Geosci. Remote, 40, 1824-1832.
32. Jonsson, P., & Eklundh, L. (2004). TIMESAT- a program for analyzing time series of satellite sensor data, Comput. Geosci., 30, 833-845.
33. Kandasamy, S., Baret, F., Verger, A., Neveux, P., & Weiss, M. (2013). A comparison of methods for smoothing and gap filling time series of remote sensing observations - application to MODIS LAI products. Biogeosciences, 10, 4055-4071.
34. Kendall, M.G. (1975). Rank Correlation Methods, 4th edition. Charles Griffin, London.
35. Kohonen T. (2001). Self-organizing maps, 3rd Ed., New York: Springer.
36. Mann, H.B. (1945). Non-parametric tests against trend. Econometrica, 13, 163-171.
37. McLeod, A.I. (2011). Kendall: Kendall rank correlation and Mann-Kendall trend test. R package version 2.2. https://CRAN.R-project.org/package=Kendall.
38. Neeti, N., & Eastman, J.R. (2011). A Contextual Mann-Kendall Approach for the Assessment of Trend Significance in Image Time Series. Transactions in GIS, 15(5). Wiley Online Library: 599-611.
39. Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. Journal of the American Statistical Association, 63(324), 1379-1389.
40. Shiskin, J., Young, A. H., & Musgrave, J. C. (1967). The X-11 variant of the Census Method II seasonal adjustment program. Technical Paper No. 15, U.S. Department of Commerce, U. S. Census Bureau.
41. Stineman, R.W. (1980). A Consistently Well Behaved Method of Interpolation. Creative Computing, 6(7), pp. 54-57.

42. Taylor, M. (2017). sinkr: Collection of functions with emphasis in multivariate data analysis. R package version 0.6. https://github.com/marchtaylor/sinkr.
43. Van den Dool, H.M., Saha, S., & Johansson, Å. (2000). Empirical orthogonal teleconnections. Journal of Climate, 13(8), 1421-1435.
44. Verbesselt, J., Hyndman, R., Newnham, G., & Culvenor, D. (2010). Detecting Trend and Seasonal Changes in Satellite Image Time Series. Remote Sensing of Environment, 114(1), 106-115.
45. Verger, A., Baret, F., & Weiss, M. (2011). A multisensor fusion approach to improve LAI time series. Remote Sens. Environ., 115, 2460-2470.
46. Verger, A., Baret, F., Weiss, M., Kandasamy, S., & Vermote, E. (2013). The CACAO method for smoothing, gap filling, and characterizing seasonal anomalies in satellite time series. *IEEE transactions on Geoscience and Remote Sensing*, *51*(4), 1963-1972.
47. Wang, D., & Liang, S. (2011). Integrating MODIS and CYCLOPES leaf area index products using empirical orthogonal functions. IEEE Transactions on Geoscience and Remote Sensing, 49(5), 1513-1519.
48. Wang, Y., & Liu, D. (2013). Reconstruction of satellite chlorophyll-a data using a modified DINEOF method: a case study in the Bohai and Yellow seas, China. Intern. Journal of Remote Sensing, v.35(1), 204-217.
49. Weiss, D.J., Atkinson, P.M., Bhatt, S., Mappin, B., Hay, S.I., & Gething, P.W. (2014). An effective approach for gap-filling continental scale remotely sensed time-series. ISPRS Journal of Photogrammetry and Remote Sensing, 98, 106–118.
50. Wilks D.S. (1995) Statistical Methods in the Atmospheric science, Academic Press. ISBN 0-12-751965-3.
51. United States Census Bureau (2016). X13-ARIMA-SEATS Reference Manual Accessible HTML Output Version 1.1.

ECOPOTENTIAL dissemination from main authors:

Filipponi F., Valentini E., Taramelli A., 2017. A collection of data analytics for Earth Observation time series analysis. Earth Observation Open Science 2017 conference, Frascati (Italy), 25-28 September 2017, oral presentation.

Filipponi F., Valentini E., Taramelli A., 2017. Sea Surface Temperature changes analysis, an Essential Climate Variable for Ecosystem Services provisioning. MultiTemp 2017 conference, Bruges (Belgium), 27-29 June 2017, oral presentation.

Valentini E., Filipponi F., Nguyen Xuan A., Passarelli F.M., Taramelli A., 2016. Marine food provision ecosystem services assessment using EO products. ESA Living Planet Symposium 2016, Prague (Czech Republic), 9-13 May 2016, poster session.