Bachelor Thesis

Detection of Changepoints in Ocean Time Series

Severin Kaderli

Supervisor: Prof. Thomas Frölicher and Friedrich Burger

Climate and Environmental Physics (CEP)
University of Bern
June 2020

Abstract

The aim of this thesis is to analyze climate time series for abrupt changes. The time series are based on data provided by simulations of the Earth System Model GFDL ESM2M and consist of a historical period and a Representative Concentration Pathway (RCP8.5) period, where high emissions of greenhouse gases are assumed. The annual means of the four ocean variables sea surface temperature, sea ice concentration, net primary production of phytoplankton, and surface partial CO_2 pressure are analyzed for abrupt changes. The analysis is based on the R package EnvCpt, which is designed to detect structural changes in climate and environment time series. Using post-processing R scripts, the EnvCpt output is filtered and presented in three different spatial plots. Several regions showing abrupt changes either caused by long-term shifts throughout the historical period or caused by changes due to the RCP8.5 period could be identified. The results are compared to the ones obtained by Drijfhout et al. (2015), who also analyzed climate time series for abrupt changes. With toy data, some limitations of EnvCpt are recognized and some recommendations for the use of EnvCpt are given.

Contents

1 Introduction

1.1 Motivation

All physical quantities describing the state of the atmosphere or the oceans are subjected to natural changes. These changes can progress slowly, but they can also exceed a threshold that causes a sudden and strong change, possibly leading to a new climate state. These abrupt changes in one variable can have a large impact on other variables, resulting in drastic consequences for ecosystems and societies [Drijfhout et al., 2015]. Therefore, it is important to analyze and understand processes that lead to such changes in order to be able to recognize them in the future. In the analysis, the definition of such abrupt changes is essential for the detection of the time when the changes occur.

Drijfhout et al. (2015) have already analyzed abrupt shifts in climate time series. They focused on model data simulated by the Coupled Model Intercomparison Project 5 (CMIP5) [Taylor et al., 2012]. The CMIP5 models include modules for the oceans, atmosphere, land surface, and cryosphere. They analyzed time series covering the historical 1861-2005 period and a future period over 2006-2100 that assumes high emissions of greenhouse gases (Representative Concentration Pathway - RCP8.5). Since they analyzed many variables on a global scale, Drijfhout et al. (2015) introduced several categories in order to distinguish and compare the occurring changes. The four categories are switch, forced transition to switch, rapid change into new state, and gradual change to new state. All of the categories additionally contain different types concerning the different ecosystems.

In order to identify regions with potential abrupt shifts, Drijfhout et al. (2015) calculated three different quantities from the annual mean of each grid cell. These quantities include the difference between the mean states of the first and the last 10 years of the time series, the standard deviation of the 10-year running mean and the maximum absolute change within the 10-year running mean. They analyzed spatial maps of these three quantities, and if two of the three quantities indicated an abrupt change, the region was studied using additional criteria. For forced shifts such a criterion is that the maximum absolute 10-year change is larger than four-times the standard deviation of the preindustrial period. Overall, they detected 37 events of forced changes in all ecosystems after a certain increase of the global temperature. The results of Drijfhout et al. (2015) are presented in more detail in the discussion section 4.1.

In this thesis, a different approach is used to identify abrupt changes in ocean time series. The fundament of the approach taken here is again the definition of points in time when the sudden changes happen. Unlike in the analysis of Drijfhout et al. (2015), the position of change is defined as the point where the statistical properties of the time series change. In order to identify these so called changepoints, different statistical models are compared with the time series. Making use of the concept of information criteria, the model delivering the best fit to the data compared to the others is determined.

This approach is implemented in the R package EnvCpt developed by R. Killick and C. Beaulieu (2018). With EnvCpt, structural changes in climate and environment time series can be detected.

1.2 Aim

As will be described in the method section 2.3, the time series used in this thesis are based on simulations produced by the Earth System Model (ESM) GFDL-ESM2M. The annual-mean time series of the four variables sea surface temperature (SST), sea ice concentration (SIC), net primary production (NPP) of phytoplankton, and surface partial CO_2 pressure (pCO_2) are analyzed. The analysis with EnvCpt has the following goals:

- The first goal is to study the overall changes in the variables, so that the change-points detected by EnvCpt can be interpreted. To do so, the evolution of the global mean, as well as the mean states of the first and the last 20 years of the time series are analyzed.

- The second aim is to identify the regions, where EnvCpt detects changes in the statistical properties of the time series. Moreover, the relation between the changes of the different variables is studied and then also compared to the results obtained by Drijfhout et al. (2015).

- Since there are several parameters which can be adjusted when using EnvCpt, we examined whether the choice of certain parameters is justified or not. Therefore the statistical concepts implemented in EnvCpt are analyzed in detail and the changes in the results that arise from parameter choices within EnvCpt and from decisions in the experimental design are studied.

- In a last step, the limitations of EnvCpt are examined. In order to do so, toy data with a specific behavior is generated and the performance of EnvCpt is analyzed.

2 Methods

2.1 Goal of the Analysis

Consider a dataset that contains a time series of a variable, as for instance sea surface temperature (SST), at different locations. Observation-based estimations and climate models projections suggest that under anthropogenic warming, sea surface temperature will increase. These changes won't be of the same kind in the Arctic Ocean as near the equator. This difference arises from a variety of processes that differ between these regions, such as the formation and melting of sea ice. The goal of the analysis is to identify changes in the properties of these time series. This is done by finding appropriate statistical models that describe these time series and by identifying changepoints, where these statisticial models undergo changes in their parameters. There are different concepts to determine if a model delivers substantial support for a time series or not. One way to do such analyses is implemented in the EnvCpt package, an R package for the detection of structural changes in climate and environment time series [Beaulieu C., Killick R., 2018].

In order to understand what the functions implemented in EnvCpt are computing, the used statistical concepts, as well as the data that is analyzed, are shortly introduced. Afterwards, the output of EnvCpt and the post-processing procedures will be explained.

2.2 Statistical Background

2.2.1 Log-Likelihood Estimation

Consider a set of independent observations $x_1, ..., x_n$ of a random variable x from a density function family $f(\cdot|\vartheta)$, where ϑ is a set of parameters used to describe the distribution. The likelihood to observe the values $x_1, ..., x_n$ is given by the likelihood function:

$$L(\vartheta) = \prod_{i=1}^{n} f(x_i|\vartheta) \tag{1}$$

The best parameter choice will result in the largest value of equation (1). Instead of considering the likelihood function directly one can use the log of the likelihood function. In that case, the log-likelihood function can be rewritten as

$$l(\vartheta) = \log L(\vartheta) = \log \prod_{i=1}^{n} f(x_i|\vartheta) = \sum_{i=1}^{n} \log f(x_i|\vartheta). \tag{2}$$

Maximizing this equation again leads to the best set of parameters $\hat{\vartheta}$ for a given density function familiy f. Assuming that the distribution of the observations $x_1, ..., x_n$ follows a normal distribution

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{3}$$

the log-likelihood function can easily be derived:

$$l(\mu, \sigma^2) = -\frac{n}{2} \log{(2\pi)} - \frac{n}{2} \log{(\sigma^2)} - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2} \tag{4}$$

The result of maximizing this equation is equivalent to getting a measure of fitness for a constant mean model fitted on a set of observations. The same procedure can be applied to fit a linear regression. Here, the assumption is made that the means of the observations y_i linearly depend on the observations x_i.

$$y_i \sim \mathcal{N}(F(x_i|\theta), \sigma^2) \tag{5}$$

with $F(x_i|\theta) = F(x_i|a, b) = a + x_i b$. This leads to the normal distribution of the residuals $y_i - F(x_i|\theta)$. This is the same assumption as made before for the mean model. Using equation (3) the log-likelihood function can be derived:

$$l(a, b, \sigma^2) = -\frac{n}{2} \log{(2\pi)} - \frac{n}{2} \log{(\sigma^2)} - \sum_{i=1}^{n} \frac{(y_i - a - b \cdot x_i)^2}{2\sigma^2}. \tag{6}$$

When using this method to fit different models to a time series, the log-likelihood function of these models, each with optimal parameter choice, sheds light on the adequacy of the models. The model with the highest log-likelihood has the lowest missmatch with respect to the time series. However, given similar log-likelihood, one should consider a model with fewer parameters to be more adequate than one with more parameters. This aspect is discussed in the next section on information.

Finally, to determine if a model is actually a good fit for the observed data, the distribution of the residuals has to be investigated. When the model delivers a substantial support, the residuals should follow the assumed distribution (here, the normal distribution is used) and they should be independent.

2.2.2 Information Criteria

As seen in the section before, the log-likelihood function can be used as a measure of how well a model fits the observations. In order to get the largest value, the adjustable parameters of the assumed distribution have to be estimated. In a model considering a constant mean the number of these parameters k is 2 and in a model based on a linear trend k is 3. Since all parameters are estimated, the more parameters are needed to be estimated the bigger the error.

In order to find the model with the best fit, the log-likelihood value has to be balanced with a penalty term counteracting the number of estimated parameters. This leads to the concept of information criteria.

Aikake Information Criterion:

The Aikake Information Criterion (AIC) is based on the log-likelihood function (2) and is defined as [H. Aikake, 1974]:

$$AIC(\hat{\vartheta}) = -2l(\hat{\vartheta}) + 2k = -2 \sum_{i=1}^{n} \log{f(x_i|\vartheta)} + 2k, \tag{7}$$

where $\hat{\vartheta}$ is the best parameter choice and k is the number of freely adjustable parameters to get $\hat{\vartheta}$. Since the AIC contains the negative maximized log-likelihood and we aim at minimizing the number of parameters k, it has to be minimized for the best fit.

Bayesian Information Criterion:

Like the AIC, the Bayesian Information Criterion (BIC) is based on the log-likelihood function (2) and is defined as [Killick et al., 2012]

$$BIC(\hat{\vartheta}) = -2l(\hat{\vartheta}) + k \cdot \log(n) = -2\sum_{i=1}^{n} \log f(x_i|\vartheta) + k \cdot \log(n), \tag{8}$$

where $\hat{\vartheta}$ is again the best parameter choice and k is the number of freely adjustable parameters to get $\hat{\vartheta}$. Unlike the AIC, the BIC depends on the number of observations n. This means that especially if n is large, a model with a lot of adjusted parameters k gets more penalized.

2.2.3   Changepoints

So far, we analyzed the fitness of models over the whole time series. However, separating the time series into different intervals and fitting the models on every interval, however may result in a better fit. This leads to the concept of *changepoints* [Killick et al., 2012]. A changepoint is a point in a time series, where the statistical properties change. Since a mean model and a trend model don't have the same statistical properties, the type of changepoint is not the same. Assume that k_{mod} is the number of parameters needed to fit a model on the whole time series. If it has m changepoints, it consists of $m+1$ pieces with differing fit parameters. So the total number of parameters used to describe the fit is

$$k = (m+1)k_{mod} + m. \tag{9}$$

Adding a lot of changepoints to a time series gets penalized in both information criteria. The next step is to find the best number of changepoints.

Optimal Partioning Algorithm:

Consider a set of observations $x_{1:n} = (x_1, ..., x_n)$ ordered in time [Killick et al., 2012]. Let's assume that the fitted model has m changepoints with the position $\tau_{1:m} = (\tau_1, ..., \tau_m)$, also ordered in time. Moreover, the first and the last observation are defined as $\tau_0 = 0$ and $\tau_m + 1 = n$. The result are $m+1$ parts of observations where the mth segment contains the observations $x_{(\tau_{-1}+1):\tau_m}$.
Like before, the aim is to minimize a function consisting of a cost function (cost for a segment) and a penalty term for the separation of the time series. The Optimal Partioning Algorithm (OP) minimizes

$$\sum_{i=1}^{m+1} \left[\mathcal{C}(x_{(\tau_{i-1}+1):\tau_i}) \right] + \beta f(m), \tag{10}$$

where $\mathcal{C}$ is a cost function and $\beta f(m)$ is a penalty term. Often, twice the negative log-likelihood is used as a cost function and the penalty term β is dependent on the number

of adjusted parameters, similarly to the information criteria. Since the cost function gets minimized by adding a lot of changepoints, the penalty term $f(m)$ is added to avoid overfitting [Haynes et al., 2015]. With the simple choice $f(m) = m$, equation (10) becomes

$$\sum_{i=1}^{m+1} \left[\mathcal{C}(x_{(\tau_{i-1}+1):\tau_i}) + \beta \right]. \tag{11}$$

The goal is now to find the partioning which minimizes equation (11) for the data set. To do so, the OP algorithm relates the optimal partitioning over the interval until the last changepoint and the segment from the last changepoint to the end of the interval. This leads to the following expression, where $F(s)$ is the minimization of equation (11) of the segment $x_{1:s}$ with changepoints $\mathcal{T}_s = \{\boldsymbol{\tau} : 0 = \tau_0 < \tau_1 < \cdots < \tau_m < \tau_{m+1} = s\}$.

$$F(s) = \min_{\boldsymbol{\tau} \in \mathcal{T}_s} \left\{ \sum_{i=1}^{m+1} \left[\mathcal{C}(x_{(\tau_{i-1}+1):\tau_i}) + \beta \right] \right\} \tag{12}$$

Setting $F(0) = -\beta$ leads to

$$\begin{aligned} F(s) &= \min_t \left\{ \min_{\boldsymbol{\tau} \in \mathcal{T}_t} \sum_{i=1}^{m} \left[\mathcal{C}(x_{(\tau_{i-1}+1):\tau_i}) + \beta \right] + \mathcal{C}(x_{(t+1):s}) + \beta \right\} \\ &= \min_t \left\{ F(t) + \mathcal{C}(x_{(t+1):s}) + \beta \right\}, \end{aligned} \tag{13}$$

where the minimal costs of the segment $x_{1:s}$ get expressed in terms of the minimal cost for the segment $x_{1:t}$ for $t < s$. From that, a recursion in s results. Starting with $F(0) = -\beta$, one loops over s from 1 to n and choses t in every step such that $F(s)$ is minimal. $F(t)$ is always known from the previous iteration steps. The computational costs are proportional to n^2.

The OP can be optimized by neglecting the points that can never be a change point, at every recursion. This algorithm is called PELT (Pruned Exact Linear Time) which has computational costs linear in the size of the time series n.

2.3   Simulated Data and Variables

The data used in this study was simulated with the Earth System Model GFDL-ESM2M, an Earth System Model (ESM) developed at NOAA's Geophysical Fluid Dynamics Laboratory (GFDL) [ESM GFDL, accessed 16.05.2020]. Within ESM2M, modules for atmosphere, ocean (-biochemistry), land, and sea ice are coupled [Dunne et al., 2012]. The ocean module MOM4p1 has a nominal horizontal resolution of 1°, which increases meridionally at the equator to $^1/_3$°. North of 65°, the grid is tripolar [Burger et al., 2020]. This leads to a total of 360 by 200 grid cells.
The simulation consists of a historical period from 1861 to 2005 and a period from 2006 to 2100 following a high-emission greenhouse gas scenario (Representative Concentration Pathway 8.5: RCP8.5) [Burger et al., 2020].
A total of four variables were analyzed at each grid cell. First, the sea surface temperature (SST) at the center of the top 10 m ocean layer. Second, the sea ice concentration

describing the fraction of the grid cell that is covered with sea ice. Third, the net primary production (NPP) of phytoplankton integrated over the top 100 m of the ocean. NPP describes how much phytoplankton nitrogen per unit time and unit area is produced. As a fourth variable the surface partial CO_2 pressure (pCO_2) is analyzed. ESM2M simulates pCO_2 pressure in the atmosphere as well as in the ocean, where in this thesis the pCO_2 of the top layer of the ocean is studied. The variables got extracted in daily steps, of which the annual means were used for the analysis.

2.4 The EnvCpt Package

The introduced statistical concepts in section 2.2 were collected in the R package EnvCpt [Beaulieu C., Killick R., 2018]. The next sub-sections describe the used models for comparison, the output of the functions implemented in the R package, and how the output gets post-processed with additional R scripts.

2.4.1 Changepoint Detection in EnvCpt

The algorithm for changepoint detection used in EnvCpt is based on the PELT algorithm with linear computational costs, described in section 2.2.3. As a penalty term, a modified version of the BIC is used [Zhang N., Siegmund D., 2007]. Like the BIC, the modified Bayesian Information Criterion (mBIC) uses the log-likelihood function. The main difference is that the penalty term differs from the one used in the BIC, since it regards the size of the data set and the spacing of the changepoints,

$$mBIC(\hat{\vartheta}) = -2l(\hat{\vartheta}) + \frac{1}{2}\left[3m\ \log(n) + \sum_{i=1}^{m+1}\log(r_i - r_{i-1})\right], \tag{14}$$

where r_i is the position of the changepoint τ_i divided by the size of the dataset n. With the mBIC, equally spaced changepoints over the observations are penalized.

2.4.2 Models in EnvCpt

EnvCpt considers basically two types of models with different modifications. In the first category are models based on a constant mean [Beaulieu C., Killick R., 2018]. So the simplest of all models consists of a constant mean μ together with a term for the noise error e_t.

$$y_t = \mu + e_t \tag{15}$$

The noise errors are assumed to be independent and identically distributed with mean value of zero and a variance of σ^2. In the second model, the equation for the first one gets modified with a first-order autocorrelation (AR1) term.

$$y_t = \mu + \varphi y_{t-1} + e_t \tag{16}$$

This term considers correlation between the observation n and $n-1$. Similarly, one can consider second-order autocorrelation (AR2), which additionally takes correlation between the observations n and $n-2$ into account.

$$y_t = \mu + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + e_t \tag{17}$$

In the second category the same considerations can be done for models based on a linear trend instead of a constant mean, leading to the following three equations:

$$y_t = \lambda + \beta t + e_t \tag{18}$$
$$y_t = \lambda + \beta t + \varphi y_{t-1} + e_t \tag{19}$$
$$y_t = \lambda + \beta t + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + e_t \tag{20}$$

As described in section 2.2.3, the time series can be separated at the points where the statistical properties change. The following equation shows the mean model fitted on a dataset separated into $m+1$ pieces (equivalent to m changepoints).

$$y_t = \begin{cases} \mu_1 + e_t, & t \leq c_1 \\ \mu_2 + e_t, & c_1 < t \leq c_2 \\ \vdots & \vdots \\ \mu_{m+1} + e_t, & c_m < t \leq n \end{cases} \tag{21}$$

The positions of the changepoints are c_m. These considerations lead to a total of 12 different models compared in EnvCpt.

Every model has its specific number of parameters k required to be adjusted for the fit. Since the variance σ is used in the noise error term, one obtains for example $k = 2$ for the mean model. Table 1 shows k for every model without changepoints.

Table 1: Number of parameters k for every model without changepoints.

mean	meanar1	meanar2	trend	trendar1	trendar2
2	3	4	3	4	5

For the models considering changepoints, k can be derived from equation (9) by using the values from Table 1 as k_{mod} and the number of changepoints m.

2.4.3 Output of EnvCpt

`envcpt()`: This is the main function of the EnvCpt package. For each of the 12 described models, the maximum of the log-likelihood function and the number of adjusted parameters are calculated. The following attributes can be adjusted in the `envcpt()` function:

- Model selection: The models considered in the execution of `envcpt()` can be adjusted. This can be useful for error handling or reducing the computation time.

- Minimal segment length: This is the minimal distance between two changepoints. The `envcpt()` function uses a default value of 5 time steps (equivalent to 5 years in our analysis).

`AIC(), BIC():` The functions for calculating the Aikake and Bayesian information criteria.

`AICweight:` Beaulieu and Killick (2018) suggest a method to weight the AIC/BIC values against each other. With the implemented `AICweight` function, the weights

$$w_i = \frac{\exp(-0.5\Delta_i)}{\sum\limits_{r=i}^{12} \exp(-0.5\Delta_r)} \tag{22}$$

can be calculated for every model, where Δ_i is given by

$$\Delta_i = AIC_i - AIC_{min}. \tag{23}$$

If only a selection of the 12 default models is considered the weights of the models changes because the denominator of the weight's definition is adjusted. The weights are always between 0 (no support) and 1 (highest support) and can therefore be interpreted as a likelihood for a model to be the most appropriate compared to the other ones. In the same way, a function for the BIC weights has been created, since it is not implemented in EnvCpt.

2.4.4 Post-Processing

Evaluation of the functions implemented in EnvCpt:

In order to execute the functions provided by EnvCpt at every grid cell, an R script has been developed. It returns a list with the following entries:

- Coordinates of the grid cell related to the underlying raster.

- Longitude and latitude of the grid cell.

- A list with informations about every considered model, containing the value of the information criterion, the difference from the best model Δ, and the related weight for both AIC and BIC.

- The positions of the changepoints of the best AIC and the best BIC model.

- A list with the changepoints detected in every model with the respective weight.

Informations about the models which are not considering changepoints are not stored in the list.

Filtering the changepoint data:

The created list consisting of the EnvCpt output at every grid cell contains a lot of informations about all detected changepoints, which needs to be filtered. There are two different R scripts for filtering either by the best AIC/BIC model or by the list containing all changepoints.

- **Filter method I:**

  The filtering method based on the best AIC/BIC model can be used to determine the type of changepoint detected in a grid cell suggested by the model with the lowest AIC/BIC value. The R script for this filter method can be found in appendix in listing 2. Besides the choice of information criterion, one can also set a minimal AIC/BIC weight needed for a changepoint model to be regarded as justified. The default value for this weight condition is 0.95. The resulting list contains only information about the grid cells fulfilling this condition.

- **Filter method II:**

  In this method, the list containing all detected changepoints is filtered. As in method I, a condition on the weight (same default value of 0.95) and the considered information criterion can be entered. Here, the weight condition also helps to reduce the large amount of changepoints with a relatively low support.

  If two models deliver similar support for a time series, their AIC/BIC weight will be in the same range. Since the sum of all weights is equal to 1 (see equation (22)), the models must share the weights and neither will probably fulfill the weight condition anymore. In order to solve this problem, the weight of the models with changepoints at similar positions are added and therefore, changepoints that are detected by more than one model, where the models together have a high weight but neither of the models alone would have, are not neglected. The interval used in this clustering process can be adjusted and has a default value of plus/minus one year. To get the most probable year of the changepoint, the iteration starts with the changepoints of the best AIC/BIC model and continues with the remaining models ordered by their AIC/BIC weights. The R script of this filtering method can be found in the appendix in listing 3.

Neglecting changepoints at the end of the time series:

According to C. Beaulieu, EnvCpt tends to detect changepoints at the end of a time series. Therefore, changepoints in a certain interval at the end were neglected. A time period of 4 time steps is set as the default option, which can be adjusted.

3 Results

3.1 Changes in the Variables

As mentioned before, the time series from 1861 to 2100 consists of the historical period and the future period following the RCP8.5 scenario. All analyzed variables undergo changes over the period, in particular during the future period. Since in some variables these changes are relatively high also relative to the fluctuations in the historical period, we expect that EnvCpt will probably detect changepoints. In order to interpret these changepoints, the overall behavior of the variables over time has to be studied.

For each variable three different plots have been created. The first plot shows the time series of the global mean from 1861 to 2100. As mentioned in section 2.3, the grid cells are not spaced evenly on the Earth's surface. The resolution of the ocean model increases especially towards the equator, resulting in smaller grid cells there. Therefore, the value of the variable in every grid cell has to be weighted with the size of the grid cell for the calculations of the global mean. The second plot shows the mean state of the period 1861-1880. This spatial plot shows the state of the variables at the start of the simulation. In the third plot the difference between the mean states of the periods 1861-1880 and 2081-2100 is shown. With this plot, the overall changes can be determined and it is especially useful for the interpretation of the changepoints regarding whether the variables increase or decrease.

In figures 1 - 4, the three plots for SST, SIC, NPP, and $p\mathrm{CO}_2$ are shown. We now discuss the main characteristics of the changes for each variable.

Sea surface temperature:

In the time series of the global mean (shown in 1(a)), the first pronounced increase of SST can be observed around 1975. After 2010, the increase follows almost a linear trend. In high latitudes SST remains approximately at preindustrial levels, as can be seen in figure 1(c). In the Southern Ocean and in the Labrador Sea south of Greenland the SST is projected to decreases, whereas in remaining regions a substantial increase is simulated.

Sea ice concentration:

The global mean of SIC shows almost the reversed curve of the global mean SST (see plot 2(a)). In 1975 a first decrease can be observed. After a short increase in the 1990s, SIC declines distinctly. Looking at plot 2(c), the largest decrease of SIC is simulated in the Arctic Ocean. In one region SIC increases, namely in the Weddell Sea southeast of Argentina.

Net primary production:

There is also a visible change in NPP at the beginning of the RCP8.5 period, but not as distinct as in the previous two variables (see plot 3(a)). The biggest changes in the NPP are projected to tropical and subtropical regions expanding up to about $\pm$ 40°, whereby the changes contain increases and decreases.
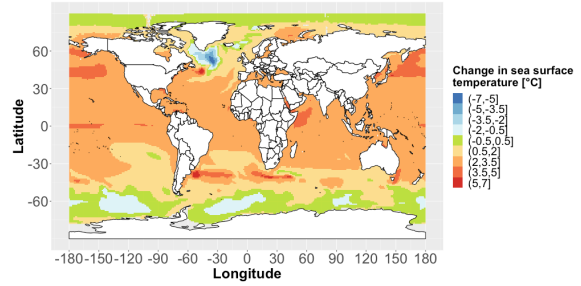
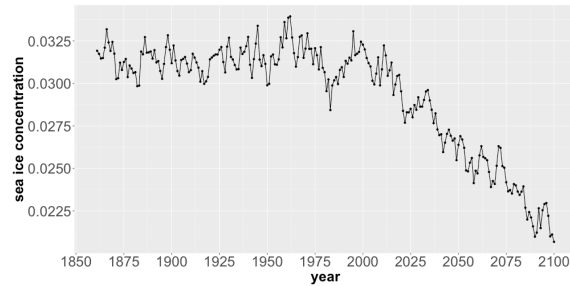# Changes in sea surface temperature
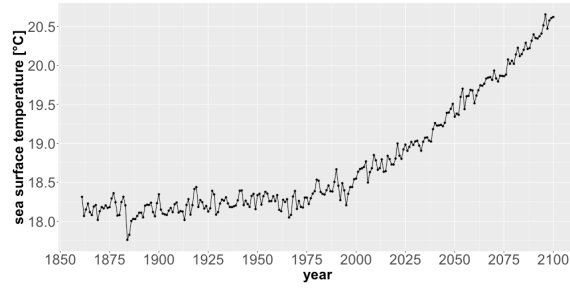
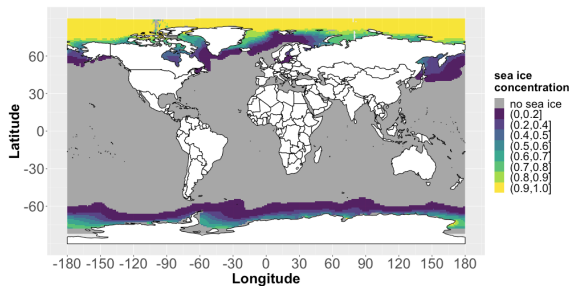(a) global mean 1861-2100
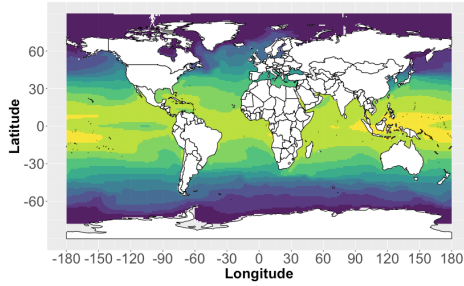
(b) mean state 1861-1880

(c) difference in the mean states between the periods 1861-1880 and 2081-2100

Figure 1: Visualization of the change of the SST; (a) global annual mean from 1861 to 2100; (b) mean state of the period 1861-1880; (c) difference between the mean states of the periods 1861-1880 and 2081-2100.
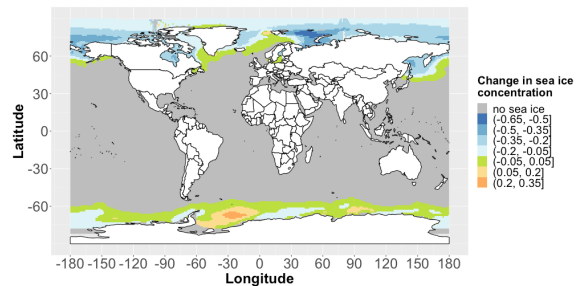
Changes in sea ice concentration
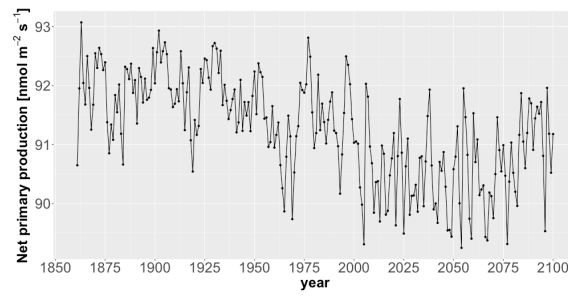

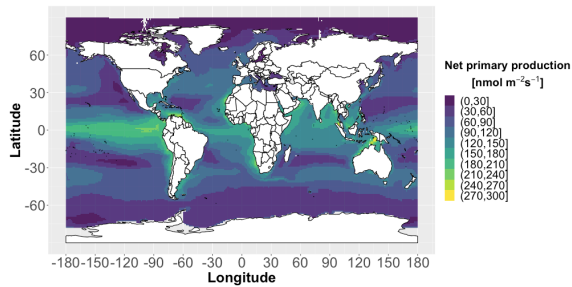(a) global mean 1861-2100

(b) mean state 1861-1880

(c) difference in the mean states between the periods 1861-1880 and 2081-2100

Figure 2: Visualization of the change of the sea ice concentration; (a) global annual mean from 1861 to 2100; (b) mean state of the period 1861-1880; (c) difference between the mean states of the periods 1861-1880 and 2081-2100.
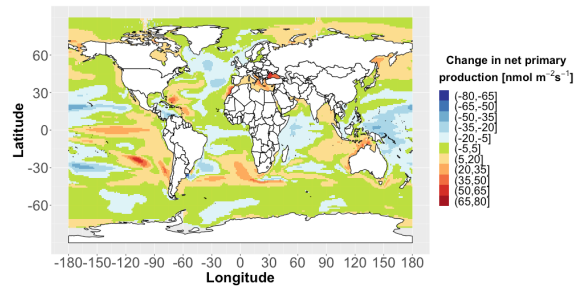
Changes in net primary production
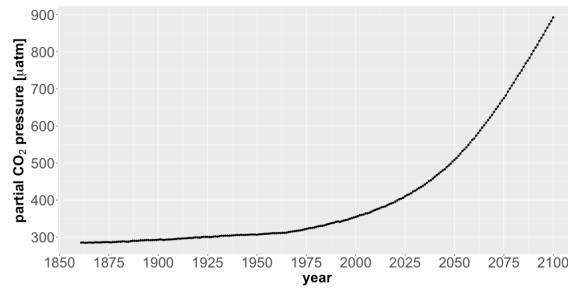

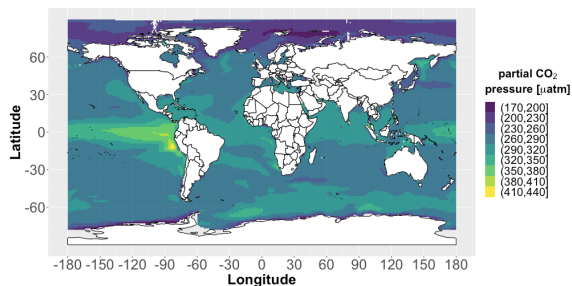(a) global mean 1861-2100

(b) mean state 1861-1880

(c) difference in the mean states between the periods 1861-1880 and 2081-2100

Figure 3: Visualization of the change of the net primary production; (a) global annual mean from 1861 to 2100; (b) mean state of the period 1861-1880; (c) difference between the mean states of the periods 1861-1880 and 2081-2100.

Changes in partial CO_2 pressure

(a) global mean 1861-2100

(b) mean state 1861-1880

(c) difference in the mean states between the periods 1861-1880 and 2081-2100

Figure 4: Visualization of the change of the partial CO_2 pressure; (a) global annual mean from 1861 to 2100; (b) mean state of the period 1861-1880; (c) difference between the mean states of the periods 1861-1880 and 2081-2100.
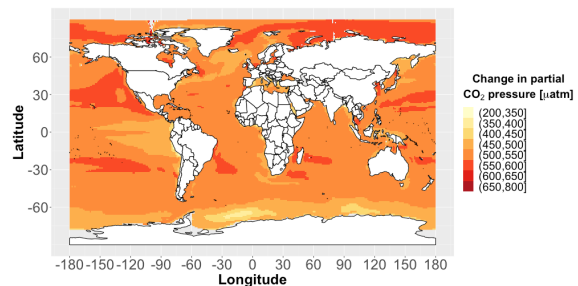
Partial CO$_2$ pressure:

The largest relative change of all variables is simulated in partial CO$_2$ pressure. From 1861 until 1975 the global mean of pCO$_2$ shown in figure 4(a) follows a nearly linear trend. After 1975, the increase is no longer linear, but follows more exponential behavior. This can also be observed in the change of the mean state in figure 4(c), where no regions with a decrease in pCO$_2$ can be found.

3.2 Visualization of Filtered EnvCpt Output

3.2.1 The Types of Plots used for the Analysis of Changepoints

Using the post-processing R scripts, the informations provided by the EnvCpt package are filtered. Depending on the different input parameters, large lists of data are resulting. In order to visualize the data, three types of plots are shown:

- **Type 1:** The first plot type shows the *total number of changepoints* detected in every grid cell. While the plot doesn't contain any information about the exact positions in time of the changepoints, regions with a high rate of changepoints can be identified. This plot is based on the filtered data where the changepoints are clustered (method II).

- **Type 2:** The second type shows the *nature of changepoints* detected in every grid cell. As mentioned in the method section (2.2.3), there are different types of changepoints. Changepoints can either separate segments of a time series with different mean values or with different trends from each other. Their segments can additionally incorporate first or second-order autocorrelation. Hence, EnvCpt considers six types of changepoints in total. This plot is based on data filtered by method I (section 2.4.4) and therefore shows which of the changepoint models has the largest support.

- **Type 3:** In a third plot type the *year of the first changepoint* is shown. For visualization, the years were grouped in decades. This plot is also based on the filtered data with clustered changepoints. Since the iteration of adding up the weights of the changepoints starts with the changepoint positions of the best AIC/BIC model, the most probable year is shown in the plot.

While such plots can be used for the spatial analysis of the post-processed EnvCpt output, they contain no information about the positions of the changepoints in general, except of the first changepoint. In order to identify regions where simultaneous changepoints occur in a large area, a plot for every time step was created. These plots are based on the filtered data with clustered changepoints and are additionally showing one of the six types of changepoint, as specified for the type 2 plot. Moreover, with these plots one can analyze if there are changepoints occurring simultaneously in more than one variable.

3.2.2 Main Characteristics of the Plots with Default Parameter Choice

In a first step, the plots based on the default parameters are analyzed. As a reminder, the default parameters are AIC information criterion, weight condition of 0.95, interval for changepoint clustering of $\pm$ 1 year, and minimal segment length of 5 years. In the following the detected changepoints of the different variables are discussed.
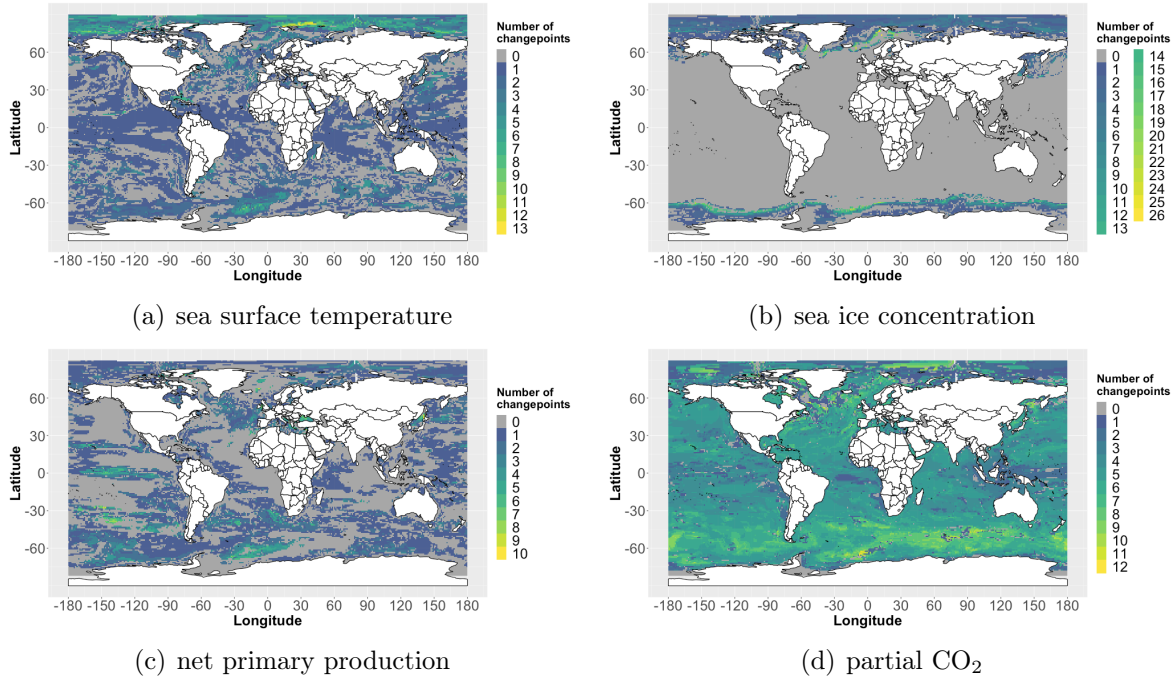
Type 1: Total number of detected changepoints

(a) sea surface temperature

(b) sea ice concentration

(c) net primary production

(d) partial CO_2

Figure 5: The number of changepoints for the four analyzed variables based on the default parameters and the time series containing the historical as well as the RCP8.5 period.

Changes in the Southern Ocean:

All variables have detected changepoints in the Southern Ocean. A hotspot of changes is located in the Weddell Sea at about 65°S and 23°E. For SST, NPP, and pCO$_2$ the total number of detected changepoints is about 10 and for SIC about 25. This can be seen in the type 1 plot of each variable in figure 5. This region also stands out in terms of the changes between 1861-1880 and 2081-2100 (figures 1-4(c)): A decrease in the SST and in the NPP and an increase in the SIC are simulated. Expect for pCO$_2$, where the changes occur in a linear trend, the nature of change in the other variables is generally not following a clear pattern. This result will be discussed later in the robustness study.

The first changepoints in all variables have been detected at the end of the 19th century (see figure 7). By analyzing the positions of all detected changepoints, large pattern with changepoints occuring in the same year could be found for all variables. Besides the common changepoints at around 1890, bigger changes at around 1990 and 2065 could be detected.

Type 2: Nature of the changepoints

(a) sea surface temperature

(b) sea ice concentration

(c) net primary production

(d) partial CO$_2$

Figure 6: The nature of changepoints for the four analyzed variables based on the default parameters and the time series containing the historical as well as the RCP8.5 period.

Type 3: Year of the first changepoint

(a) sea surface temperature

(b) sea ice concentration

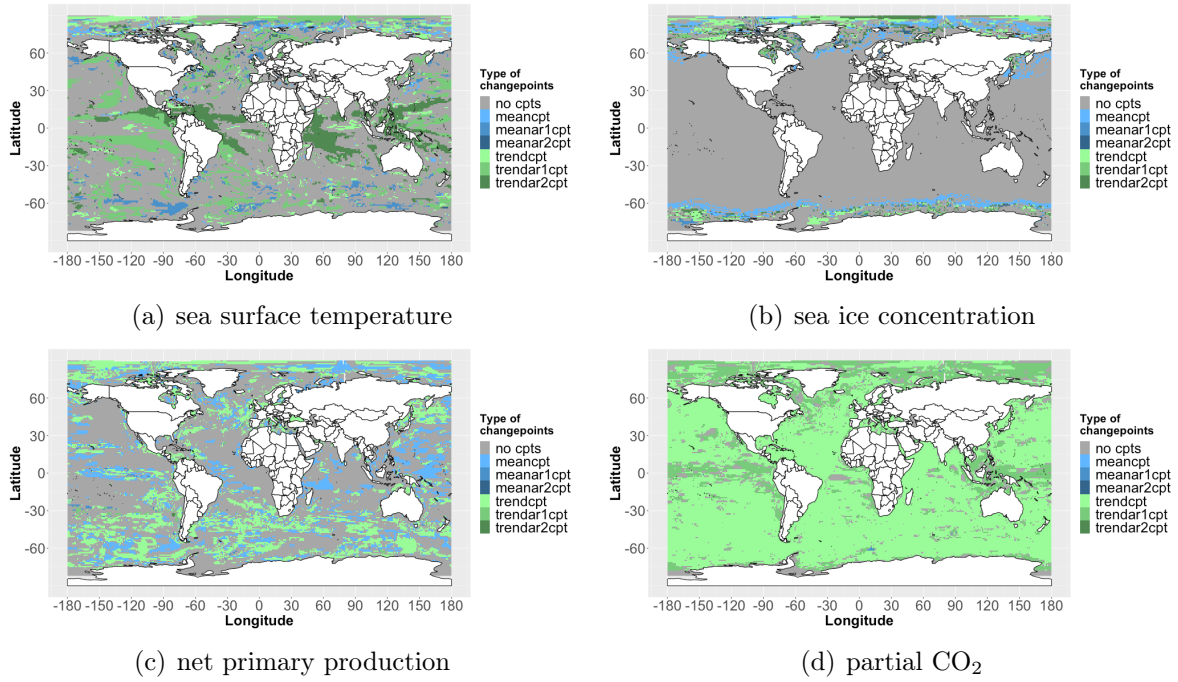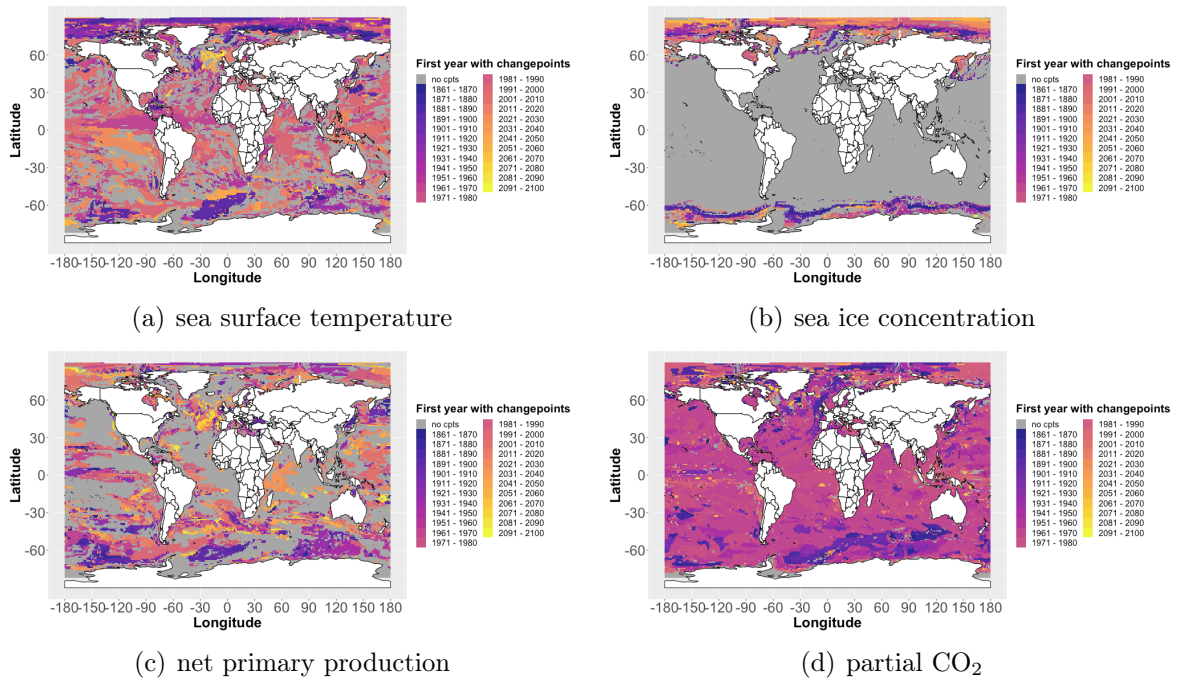(c) net primary production

(d) partial CO$_2$

Figure 7: Year of first changepoint for the four analyzed variables based on the default parameters and the time series containing the historical as well as the RCP8.5 period.

Changes in the Arctic Ocean:

Another region with a large amount of detected changepoints is the Arctic Ocean. From figure 5 we see that this is particularly true for SST which shows large patterns with more than 10 detected changepoints, while less changepoints are observed for SIC and NPP. By looking at figure 1(c) one can see that the SST in the high latitudes of the Arctic Ocean does not change much. A possible explanation is that SST shows fluctuations, which are already present in the preindustrial period leading to a large amount of changepoints.

Moreover, in the type 2 plot for SST in figure 6(a) it can be seen that the changepoints detected near the poles are often of the type trend followed by changepoints in the mean going towards the polar circles. This can also be observed for other variables like SIC and in other regions, as for example the Southern Ocean. A possible explanation for these changes in the mean is that there are shifts between periods with more and periods with less sea ice. This statement is supported by figure 7(b) where one can see that the first changepoints in these regions with mean shift occur early in the historical period.

Moreover it can be seen that regions in higher-latitudes close to the pole show their first changepoint in the middle of the RCP8.5 period for both SST and SIC. Considering the overall changes in SST and in SIC and furthermore taking into account that these changepoints are of the type trend suggest that there is a decrease in the sea ice starting at around 2030.

The same can also be observed in the Southern Ocean. As shown in plot 2(b), the SIC at around 70°N and 70°S is lower than in regions further polewards. Through the rise in SST, the little ice present at around 70 degrees North is melting completely and since after that the SIC is already zero and no more changes, except of the new formation of sea ice, can occur, a changepoint in the mean is resulting. Since the sea ice further polewards is not completely melted and the melting process can happen over the whole time series, EnvCpt interprets the change as a linear trend. Figure 8 shows the time series of SIC at coordinates 83.41°N and 2.54°E and the result of EnvCpt including the AIC/BIC value. This grid cell is close to North Pole and troughout the historical period permantely coverd by sea ice.
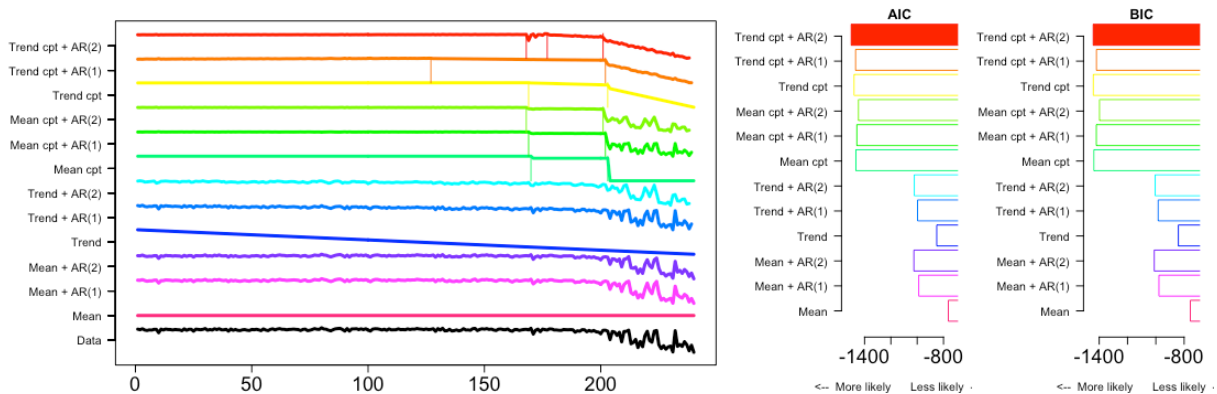

Figure 8: Plot of the time series of the SIC at the coordinates 88.46°N and 2.54°E with the output of EnvCpt including the AIC/BIC value. The filled bar indicates the best AIC/BIC model.

Following the black curve at the bottom of figure 8, a decrease in the SIC starting at time step 200 (year 2061) and another detected changepoint at time step 170 (year 2031) can be observed, already mentioned above. With both information criteria, EnvCpt interprets the changes as a linear trend with second-order autocorrelation. From that one can see why EnvCpt detects changes in trend in high-latitudes. The change in the mean in slightly lower-latitude regions is discussed in section 4.1.

Again, the plots showing only the changepoints detected in one year can be analyzed in order to see if the changes can be observed in all of the variables simultaneously. In the years around 1980, 2020 and 2065 larger patterns of changepoints could be detected in SST and in SIC.

Changes in the region south of Iceland:

Using the plots of type 3 in figure 7, a region with their first detected changepoint only in the second half of the 21st century can be identified. Located south of Iceland, the region particularly stands out when analyzing the SST and the NPP. Since the changepoints are first occurring towards the end of the time series, the total number of changepoints is not very high and therefore not notable only considering the type 1 plot. By looking at plot 1(c) one can see that SST decreases over the whole time series. This decline in water temperature and the related decline in NPP is likely related to the weakening of the meridional overturning circulation [Drijfhout et al., 2012].

Changes in low-latitude regions:

By looking at figure 7(a), showing the year of the first changepoint for SST, one can see that in low-latitude regions the first changepoints often occur in the period of 1970-2000. Again considering the overall changes in SST in figure 1(a), the global mean increase starts in the same period. Since the low-latitude regions make up a large part of the Earth's surface, they have a large influence on the global mean and it is therefore no surprise that their behavior is similar.

Changes in the partial CO_2 pressure:

In figure 6(d) the type of changepoint for pCO_2 is shown. It can be clearly seen that almost in every grid cell the suggested model is considering piecewise linear trends that are separated by changepoints. The reason for that is based on the large changes in pCO_2. It is discussed in detail later in the discussion section (4.2).

3.3   Robustness of the Results

The plots presented in section 3.2.2 depend on the choice of input parameters. Since changing only one parameter already can have a significant effect on the results, the differences have to be studied carefully.

3.3.1 Varying the Information Criterion

By looking at the definitions of the information criteria in equation (7) and (8) one can see that the only difference between the information criteria arises from the penalty terms. More precisely, the factor in front of k, the number of adjustable parameters for the fit, differs and in case of the BIC additionally depends on the number of observations n. Table 2 shows the penalty term for the full 240-year period and for the RCP8.5 period only.

Table 2: Penalty term for the different time series.

	AIC	BIC
RCP8.5	$2k$	$4.55k$
historical + RCP8.5	$2k$	$5.48k$

If BIC is used, models with a large k are penalized more. From equation (9) we see that a large number of changepoints m leads to a large k. The number of changepoints is independent of the choice of information criterion used for getting a measure of fitness. Therefore, k_{mod} and thus k are the same for the analysis with AIC and BIC.

Although the number of changepoints for each changepoint model is the same, the plots of type 1 are different (compare figure 5(a) and 21(a) in the appendix). This arises from the fact that the type 1 plot is based on the filtered data with clustered changepoints. A different value of the information criterion leads to a different weight, and since the plots are only showing changepoints according to a certain weight condition, the result may differ. Larger changes than in the plots showing the number of changepoints can be seen by comparing the type 2 plots, shown in figure 9. We observe two main differences:

(a) AIC (b) BIC
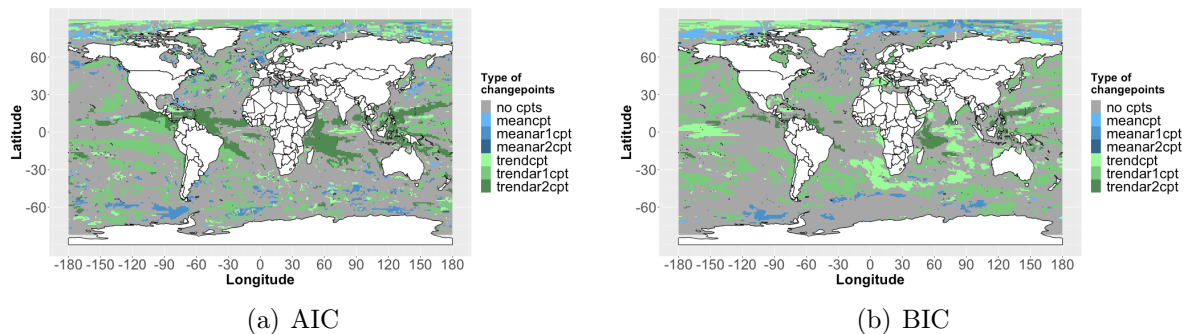
Figure 9: Type 2 plot for SST with default parameters except for the information criterion.

- **The best BIC model is not as often a changepoint model:**
  If the number of changepoints m is large, the models with changepoints are penalized more when the BIC is used. For the BIC to suggest a non-changepoint model, meaning that the penalty is large enough so that another model gets the highest support, there has to be at least one changepoint model and one without changepoints with similar log-likelihoods. If this is not fulfilled, the best BIC model is still considering changepoints.

This can for example be observed in the Weddell Sea, e.g. at 64.5°S, 22.5°W, or in the region south of Iceland, where a lot of changepoints are detected when using the AIC. Figure 10 shows the changes in the time series in SST at the coordinates mentioned above in the Weddell Sea.
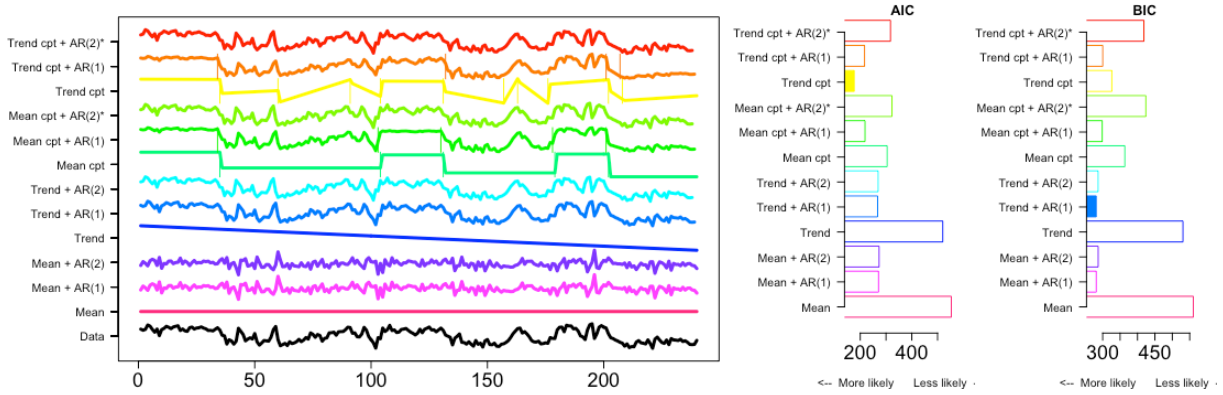

Figure 10: Plot of the time series of the SST at the coordinates 64.5°S and 22.5°W with the output of EnvCpt including the AIC/BIC value. The filled bar indicates the best AIC/BIC model.

The best AIC model in figure 10 is a trend model with changepoints. Since the number of detected changepoints in every model is quite large (from 6 to 11), they are more penalized using the BIC. Moreover, there are models with and without changepoints with a log-likelihood in the same range. As a result, the best BIC model is a model without changepoints.

- **The BIC detects new changepoints:**
  As described above, the type 2 plot is based on the data filtered by the best AIC/BIC model (filter method I). Since the default parameters contain a weight condition, only grid cells where the best model has a weight higher than 0.95 are shown. Considering the definition of the weight in equation (22), the sum of all weights of the models for a time series has to be equal to 1. Therefore, if two models have a similar AIC or BIC, neither of the models has a weight high enough to be considered. This is the explanation why there are changepoints only detected by the BIC. The larger penalty term caused by a large number of detected changpoints increases the BIC of some models sufficiently enough so that one model with a weight over 0.95 remains. By comparing plot 9(b) to the plots in figure 13, one can see that larger patterns detected by the BIC are also shown in the figures based on AIC with a weight condition of 0.5, as for example in the equator regions of the Pacific Ocean or in the region southwest of Africa.

Overall, the detected patterns stay similar when changing the information criterion in many regions. This can for example be observed in high-latitude regions. In contrast to the two differences in the results that can occur, the following criteria must be met to produce the same result. There has to be a model which provides much more support than any other model and therefore claims almost all of the weight. If the difference in the log-likelihood is large enough so that even the larger penalty of the BIC does not have an influence on the distribution of the weights, the result of AIC and BIC will be the same.

3.3.2 Varying the Minimal Segment Length

The minimal segment length is used by the OP algorithm (section 2.2.3) and it defines the minimal period between two changepoints. The default value used in EnvCpt is 5 years. In order to analyze the impact of changing this parameter, the default plots with a segment length of 10 years have been created. Changes only occur if the data based on a minimal segment length of 5 years contains changepoints separated by less than 10 years. By comparing the plots of type 1 (figure 11) one can see that there are only minor changes. Hence, the majority of changepoints are separated by more than 10 years.
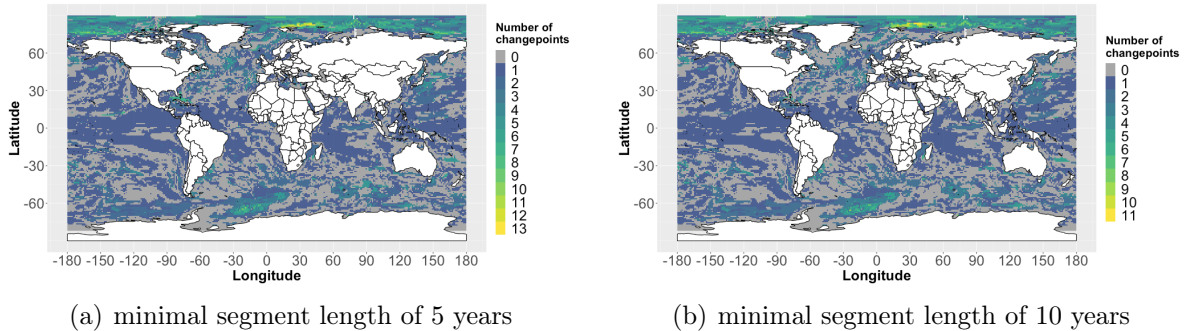

(a) minimal segment length of 5 years   (b) minimal segment length of 10 years

Figure 11: Type 1 plot for the SST with default parameters and modified minimal segment length: (a) segment length of 5 years; (b) segment length of 10 years.

3.3.3 Varying the Length of the Time Series

Changing the length of the considered time period can have a major impact on the results. This can be seen especially by comparing the plots of type 2 of the time period consisting of the historical and the RCP8.5 part and the period only consisting of the RCP8.5 part. Both plots are shown in the bottom row of figure 12. There are two possible reasons for these differences:

1. The changepoint detection is based on the OP algorithm. As described in section 2.2.3, the algorithm separates the time series in segments $x_{1:s}$ and minimizes equation (13) by iteration. By changing the length of the time series, different changepoints may lead to the minimization of equation (13). Therefore, the changepoints may not be the same, even in the part of the period, where both time series are based on RCP8.5.

2. Longer time series can contain more changepoints. Following the considerations above in section 3.3.1, the number of changepoints can have a large impact on the penalty term and therefore on choice of the model with the highest support. In the analysis based on the time period containing also the historical part, the penalty term is higher because there are often changepoints detected before 2006. As mentioned above, this leads to models without changepoints to be preferred over models with changepoints, especially when they have a similar log-likelihood.
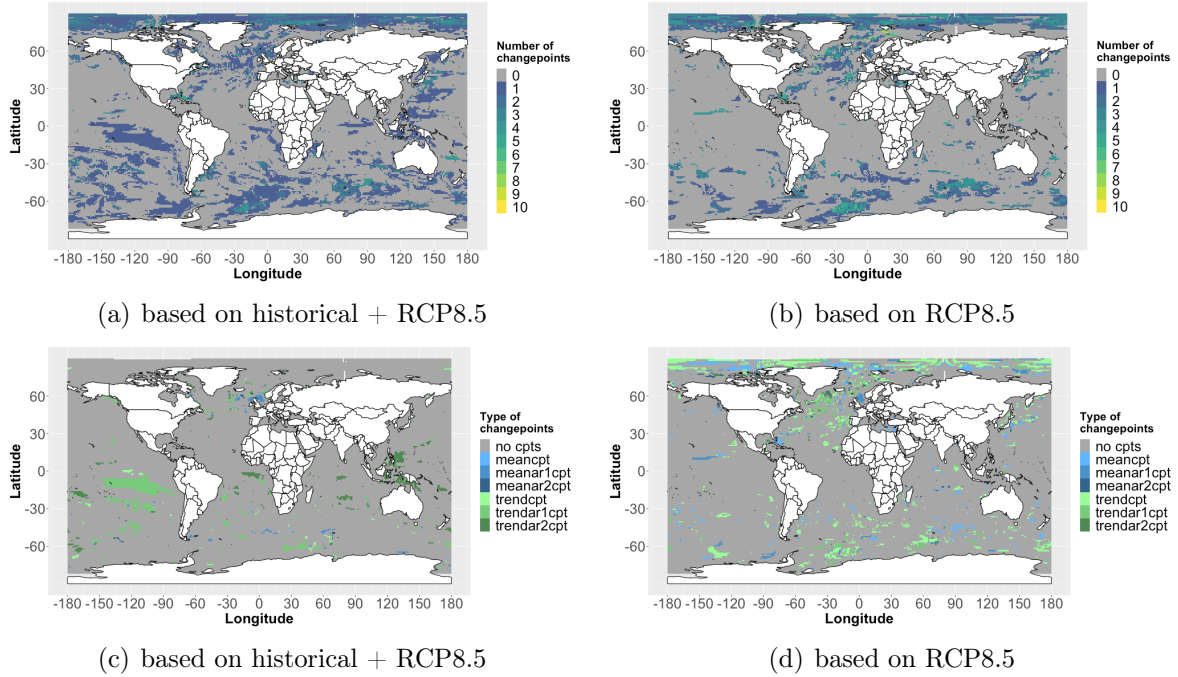
21

(a) based on historical + RCP8.5

(b) based on RCP8.5

(c) based on historical + RCP8.5

(d) based on RCP8.5

Figure 12: Top row: Type 1 plot for SST with default parameters for changepoints occurring in the years from 2006 to 2100: (a) based on the period historical + RCP8.5; (b) based only on the RCP8.5 period. Bottom row: Type 2 plot for SST with default parameters for changepoints occurring in the years from 2006 to 2100: (c) based on the period historical + RCP8.5; (d) based only on the RCP8.5 period.

The largest differences between the type 2 plots can be observed in the Arctic Ocean. While EnvCpt detects a lot of changepoints in this region when only the RCP8.5 period is considered, the plot based on both periods showing only the changepoints after 2005 indicates almost no changes. These differences can be likely explained by the data on which the type 2 plot is based on (see 2.4.4, filter method I), as well as the second argument mentioned above. The type 2 plot including the historical and the RCP8.5 period, shows the best AIC/BIC model which are penalized with the total number of changepoints found in the whole time series. As one can see in the type 1 plot shown in the top row of figure 12, the number of detected changepoints in many regions of the Arctic Ocean is the same for both time series. Therefore, following the second argument, there have to be a relatively large amount of changepoints detected in the historical period in this region leading to a large penalty term, which in turn leads to a non-changepoint model delivering the highest support. The type 1 plot for SST in figure 5(a) is indicating a high rate of changepoints in the Arctic Ocean, consistent with an early first changepoint shown in the type 3 plot in figure 7(a).

Moreover, a large pattern in the low latitudes of the Pacific Ocean can be observed in the plots based on both periods, which is not occurring in the plots only based on RCP8.5. This can likely be explained by the large increase of the SST in the RCP8.5 period. The changepoints in this region are probably caused by the change in the statistical properties in the time series based on both periods causing changepoints which are not observable when only the RCP8.5 period is considered. Again, figure 7(a) showing the year of the first changepoint for SST supports this explanation.

22

3.3.4 Varying the Weight Condition

The weight condition defines the minimal weight a changepoint model (in filter method I, see section 2.4.4) or a changepoint (in filter method II) must have to be considered. The type 1 plot with default parameters but different weight conditions can be found in figure 13. At a minimum weight of 0.5 almost every grid cell contains changepoints. When the minimum weight is increased, the fraction of grid cells with changepoints is reduced. However, there are still many grid cells with change points when choosing a high minimum weight such as 0.99.
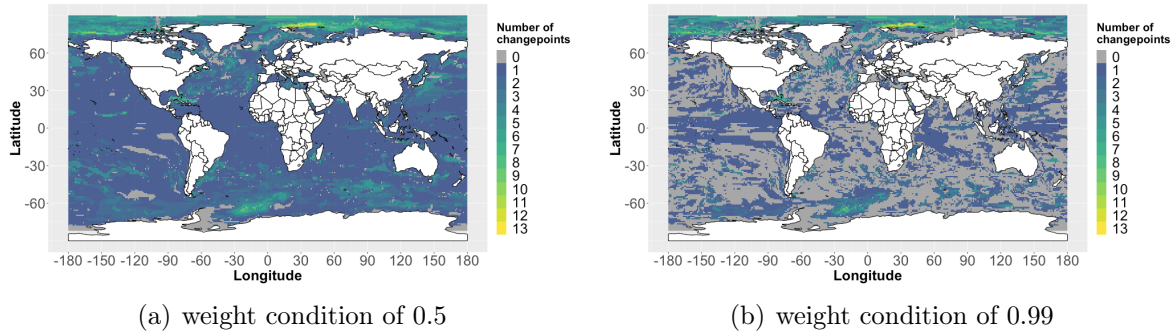

(a) weight condition of 0.5   (b) weight condition of 0.99

Figure 13: Type 1 plot for the SST with default parameters and modified weight condition: (a) weight condition of 0.5; (b) weight condition of 0.99.

3.3.5 Varying the Interval for Changepoint Clustering

The default interval for changepoint clustering is 1, meaning that the weight of the changepoints ±1 year away from the changepoint of the model with the highest AIC/BIC are added up. Increasing this interval to ±2 years results in the detection of a few more changepoints. This happens because more weight is added to the same changepoint. If the added up weight then fulfills the weight condition, additional changepoints are detected. This can be seen in figure 14 where the interval is increased to ± 2 years. Since there are only a few changes, adding up changepoints in the interval ±1 year is sufficient. With a larger interval the precision of the detection gets worse.

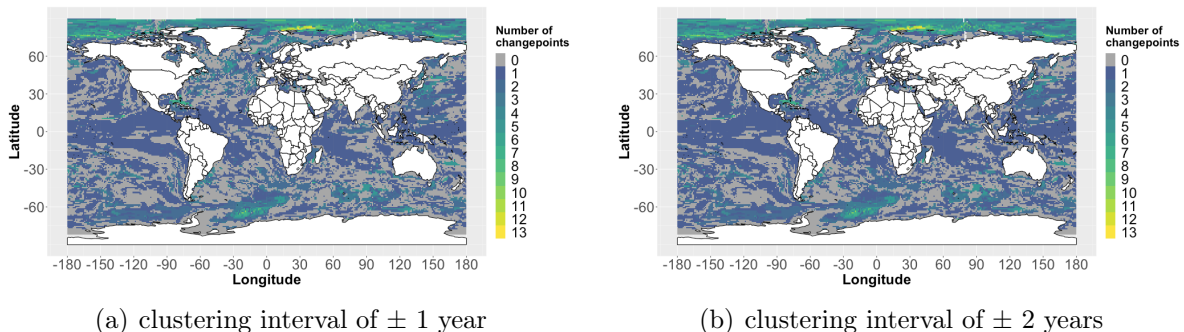(a) clustering interval of ± 1 year   (b) clustering interval of ± 2 years

Figure 14: Type 1 plot for the SST with default parameters and modified interval for changepoints clustering: (a) interval of ± 1 year; (b) interval of ± 2 years.

4 Discussion

4.1 Comparing Results to Drijfhout et al.

As mentioned in the introduction, Drijfhout et al. (2015) separated the detected abrupt shifts in four categories, whereby these categories are further subdivided by region and by defining types for different variables. Table 3 shows a selection of categories and types, which are compared to the results obtained with EnvCpt.

Table 3: Selection of categories and types of shifts detected by Drijfhout et al. (2015), which are compared to the result of the analysis with EnvCpt.

Category	Type	Region
I: switch	1. sea ice bimodality	Southern Ocean
II: forced transition to switch	2. sea ice bimodality	Southern Ocean
III: rapid change to new state	4. winter sea ice collapse	Arctic Ocean
	5. abrupt decrease in sea ice	regions of high-latitude oceans

Category I & II - Types 1. and 2.:

Drijfhout et al. (2015) identified four regions in the Southern Ocean showing reversing abrupt shifts in the sea ice cover. These changes could be observed in the historical period as well as in the RCP scenarios, whereby in the latter the abrupt shifts were becoming weaker. They explained these bimodal shifts by the feedback between the open-ocean convection and the formation of sea ice. During periods when warm water is upwelling, no permanent sea ice cover is formed. In contrast, once a permanent ice cover has been formed, the convection decreases, which enhances again the formation of sea ice.

Two of the four regions with bimodal switches identified by Drijfhout et al. (2015) are located in the Weddell Sea. Looking at figure 5(b), the same region is showing a large rate of changepoints. Often, changes between mean states are simulated, which extends to the Antarctic Polar Front, as can be seen in figure 6(b).

These observation lead to the possible conclusion that SIC is switching between different states in our analysis, whereby according to Drijfhout et al. (2015) longer periods without sea ice cover can occur. In order to check this statement, a representative grid cell in the Weddell Sea with a high rate of changepoints in the mean is analyzed. Figure 15 shows the time series of SIC for that grid cell.

In the time series at the bottom of figure 15 one can see that longer periods with a very low sea ice concentration are interrupted by periods with an increased fraction of sea ice. EnvCpt interprets the changes in the time series as changes between different mean states and therefore delivers similar results as obtained by Drijfhout et al. (2015).

Drijfhout et al. (2015) also found bimodal shifts that occur only after a change in the climate. Such regions are indicated by a relatively late first changepoint. Looking at figure 7(b) a region in the Ross Sea around 75°S and 160°W can be identified with changepoints
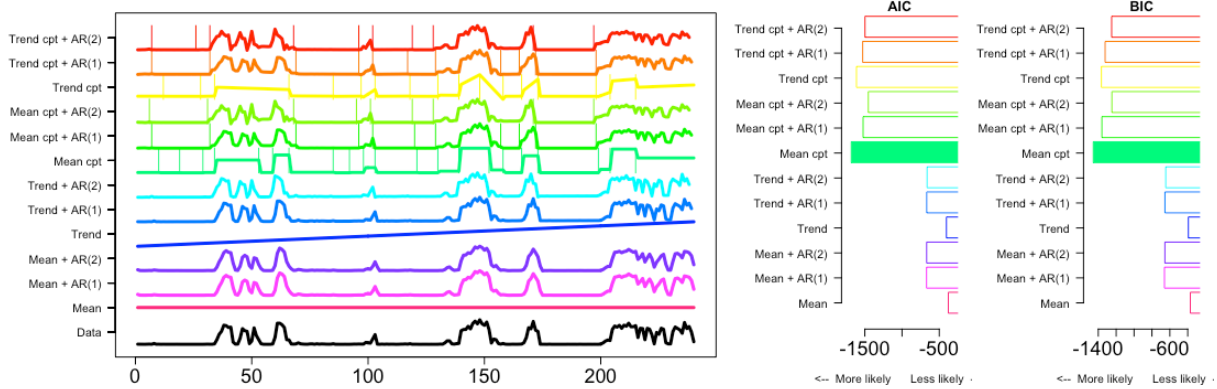
Figure 15: Plot of the time series of the SIC at the coordinates 65.5°S and 13.5°W in the Weddell Sea with the output of EnvCpt including the AIC/BIC value. The filled bar indicates the best AIC/BIC model.

first occurring around the year 2060. Moreover are these late changepoints mostly of the type mean (see figure 6(b)). Combining both observations leads to the conclusion that the changes caused by the RCP8.5 period are leading to a switch in the sea ice concentration. Figure 16 shows the time series in SIC of a grid cell located at 75.5°S and 160.5°W in the Southern Ocean.
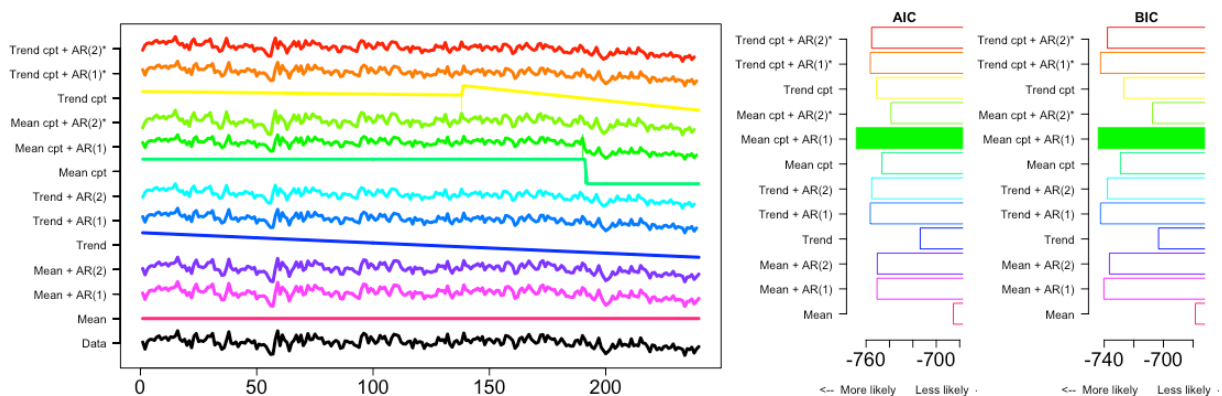

Figure 16: Plot of the time series of the SIC at the coordinates 75.5°S and 160.5°W in the Southern Ocean with the output of EnvCpt including the AIC/BIC value. The filled bar indicates the best AIC/BIC model.

The time series at the bottom of figure 16 shows a large variability until time step 200 interpreted by EnvCpt as a first-order autocorrelation process followed by decrease in the SIC interpreted as a switch in the mean. The detected changepoint occurs in the year 2051, a few years before the visible change in the time series. Drijfhout et al. (2015) found changepoints in the same region, besides shifts of category 3 type 5 discussed in the next paragraph.

Category III - Type 4:

In the time series considering RCP8.5, Drijfhout et al. (2015) detected an abrupt large-scale collapse of the winter sea ice in the Arctic Ocean after the year 2100. This collapse occur in regions where the water temperature in the summer months gets sufficiently high, so that no new ice layer will form in the winter. In this thesis only time series until the

25

year 2100 are considered. Nevertheless, it can be analyzed whether in our analysis a sea ice collapse before 2100 in the Arctic Ocean can be observed or if the regions identified by Drijfhout et al. (2015) indicate the probable collapse.

Drijfhout et al. (2015) found 5 cases of large-scale collapses in winter sea ice in the High Arctic. While figure 5(b) gives no clear indications for large changes, in the type 2 plot for SIC in figure 6(b) larger patterns of trend changepoints can be identified. In section 3.2.2 a grid cell near the North Pole at coordinates 83.41°N and 2.54°E was analyzed (see also figure 8). As observed before, SIC is close to 1 throughout the historical period and after time step 200 (year 2061) a strong decrease can be observed. At the end of the time series in the year 2100 the SIC has a value of 0.75. While in this grid cell no collapse appears, a strong decreasing trend in SIC can be identified also indicated by EnvCpt using both information criteria. Moreover, in grid cells located south of the one analyzed in figure 8, no collapse of the sea ice cover such as those identified by Drijfhout et al. (2015) could be found.

Category III - Type 5:

Another type of abrupt change observed by Drijfhout et al. (2015) is the local sea ice loss. In two cases they observed a fast transition from perennial ice cover to a state, where the ice cover almost disappears throughout the year. These cases are located in the Barents Sea and in the Pacific sector of the Southern Ocean. These local changes in the sea ice concentration also have an influence on the net primary production. In general, they observed a connection between the decrease of the sea ice cover and the rapid increase of the net primary production. This is most likely due to the increase of available light in the top layer of the ocean.

Looking at figure 5(c) one can see that in the Barents Sea indeed a region with a higher rate of changepoints can be identified. Also in the Pacific sector of the Southern Ocean an area with a lot changes in the NPP can be observed, but since the changes in the Barents Sea can be better related to the changes in SIC (see figure 5(b)) we will focus on this region by analyzing a representative grid cell. The two plots in figure 17 are showing time series of NPP and SIC of a grid cell located at 76.13°N and 40.37°E in the Barents Sea.

Comparing the time series of NPP (top) and SIC (bottom) in figure 17 one can see that the net primary production highly anticorrelates with sea ice concentration (Pearson correlation coefficient $r = -0.9339$). Except for a few peaks, the SIC decreases significantly after time step 150 (year 2011), which is resulting in a permanent increase in NPP. The fast transitions from perennial sea ice cover to a state where no more sea ice is present, which were detected by Drijfhout et al. (2015), can also be observed. After the year 2051 the sea ice cover has almost disappeared completely. Since in most time steps after 2051 the SIC is exactly zero, the grid cell is annually ice-free.
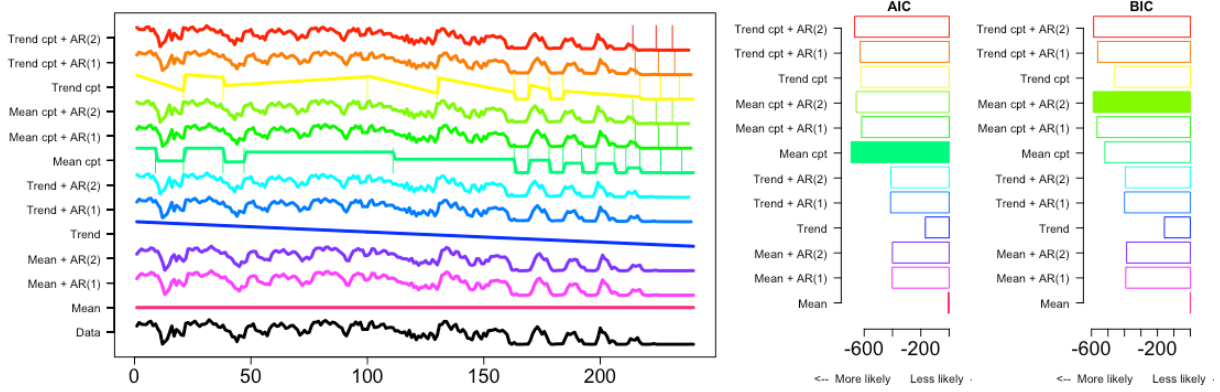
(a) sea ice concentration
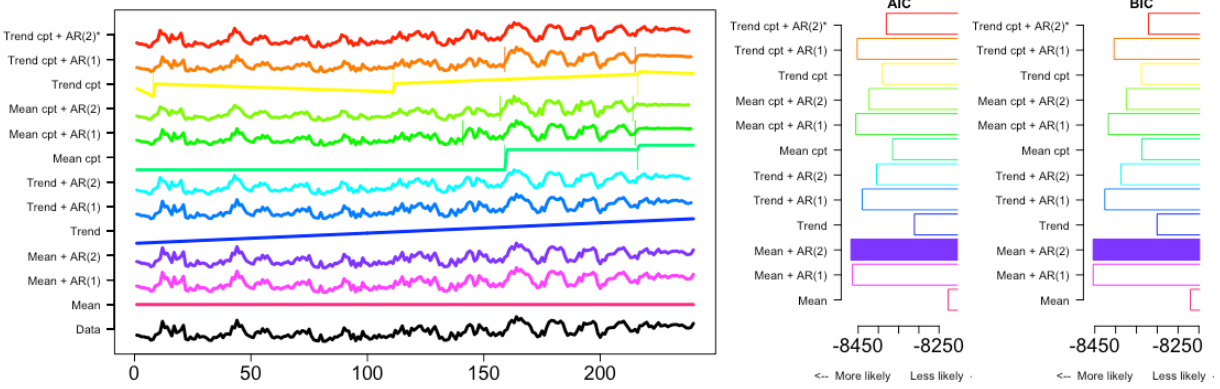
(b) net primary production

Figure 17: Plot of the time series at coordinates 76.13°N and 40.37°E in the Barents Sea: (a) for SIC; (b) for NPP. Both plots also show the output of EnvCpt including the AIC/BIC value. The filled bar indicates the best AIC/BIC model.

4.2 Performance of EnvCpt using Simulated Toy Data

EnvCpt takes into account in general two different types of models, one kind is based on a constant mean and the other on a linear trend. How does EnvCpt perform if there is an increase of higher order? In order to analyze such and similar problems the output of EnvCpt has been studied using toy (i.e. synthetic) data. With this also some limitations of EnvCpt can be recognized.

The toy data used for the analysis is generated in R using the `arima.sim()` function. This function can be used to simulate noise error as well as autocorrelation. In order to simulate autocorrelation one has to specify the autocorrelation-coefficients φ_i (see for example equation (17)). They are chosen in a way that the sum of all coefficients is equal to 0.5 in order to simulate realistic behavior. In order to get an overall behavior like a linear trend or a quadratic increase, a polynomial function of the respective degree is added. The R code used for the toy data analysis can be found in the appendix in listings 4 and 5.

In the following, three cases of special time series are studied. These examples show how EnvCpt performs on toy data with predefined behavior and where the limitations of in changepoint analysis with EnvCpt are.

27

4.2.1 Mean Model with one Changepoint

In this first analysis a time series over 240 time steps with a shift in the mean has been simulated. To this time series normal distributed noise error with a mean of zero and a standard deviation of one was added. The shift in the mean is implemented at position 120 and the mean shifts by four standard deviations. The black curve at the bottom of figure 18 shows the time series.
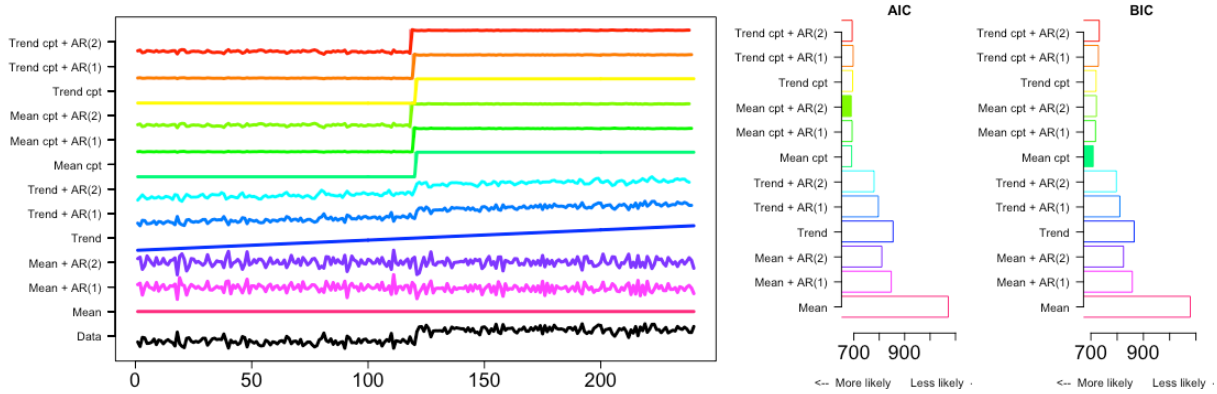

Figure 18: Result of EnvCpt for a simulated time series based on a mean with a changepoint at position 120 and added noise error, which is normal distributed with a mean of zero and a standard deviation of one. The shift in the mean is 4 standard deviations. The filled bar indicates the best AIC/BIC model.

By looking at the curves of the models produced by EnvCpt one notices that all of the models considering changepoints have detected the change at the right position. Moreover, the trend models describe the time series also quite well, what can be seen in the AIC/BIC of the respective models. Table 4 shows the log-likelihood, the number of estimated parameters, the number of detected changepoints, and the AIC/BIC for all the models.

Table 4: Log-likelihood, number of estimated parameters, number of changepoints, AIC/BIC and the AIC/BIC delivered delivered by EnvCpt for the time series shown in figure 18. Highlighted is the lowest AIC/BIC resp. the best suggested model by EnvCpt.

	mean	meancpt	meanar1	meanar2	meanar1cpt	meanar2cpt
log-like	1067.8	681.5	841.2	802.3	679.3	671.8
nr. of par.	2	5	3	4	7	9
nr. of cpts	0	1	0	0	1	1
AIC	1071.8	691.5	847.2	810.3	693.3	689.8
AIC weight	0	0.236	0	0	0.093	0.547
BIC	1078.8	708.9	857.7	824.2	717.7	721.1
BIC weight	0	0.982	0	0	0.012	0.002

	trend	trendcpt	trendar1	trendar2	trendar1cpt	trendar2cpt
log-like	848.5	681.4	788.6	769.7	679.3	671.6
nr. of par.	3	7	4	5	9	11
nr. of cpts	0	1	0	0	1	1
AIC	854.5	695.4	796.6	779.7	697.3	693.6
AIC weight	0	0.032	0	0	0.013	0.080
BIC	864.9	719.8	810.5	797.1	728.6	731.9
BIC weight	0	0.004	0	0	0	0

There is no big difference between the log-likelihood of the models considering a mean and the respective model based on a linear trend. The slope of the linear increase in the trend models is close to zero leading to similar log-likelihood values. However, due to the different number of estimated parameters, the trend models are penalized more. If EnvCpt would neglect parameters whose estimated value is close to zero, the AIC/BIC would be equal to one of the mean models.

But since the AIC/BIC of the different models are still in close range, the weights of the models are also similar. Therefore, the weight of the changepoint only considering the best AIC/BIC model may not be higher than the default weight condition of 0.95. So even if the time series is simulated with a distinct changepoint, EnvCpt will not detect the change under the weight condition. Therefore, adding up the weight of the different models in case when they detect a similar changepoint is a reasonable approach to solve this problem.

Moreover, it can be observed that the AIC often tends to interpret the noise error as autocorrelation. Using the BIC this does happen less frequently, since the penalty term of the BIC is larger.

4.2.2 Quadratic Increase

The curve of the global mean of the partial CO_2 pressure in plot 4(a) shows an increase of a higher order than linear. As a result, EnvCpt is detecting a lot of changepoints in the models considering a linear trend, as can be seen in plot 6(d).

In order to analyze this result, toy data with 240 time steps considering a quadratic increase has been simulated. Again, normal distributed noise error with a mean of zero and a standard deviation of one was added to the quadratic behavior. The total increase between the end and the beginning of the time series is equal to 50 standard deviations. The time series is shown at the bottom of figure 19.
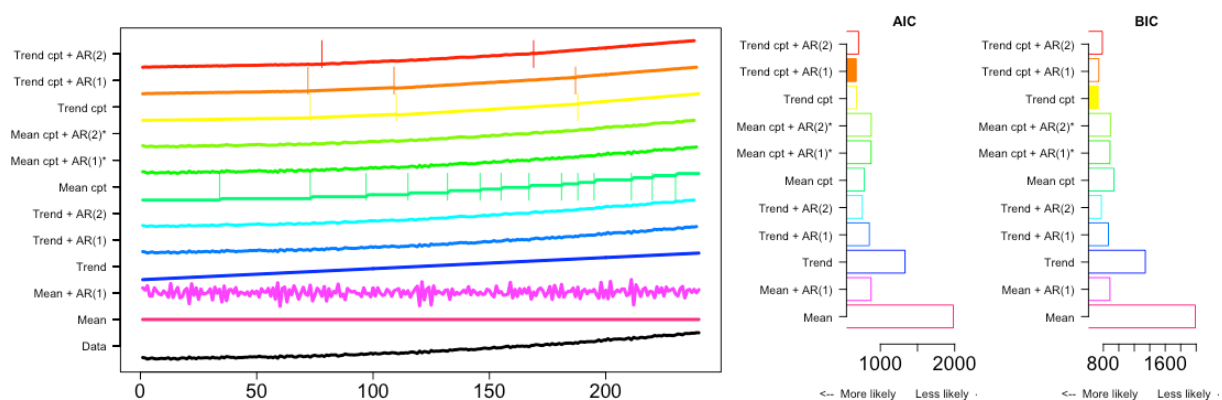

Figure 19: Result of EnvCpt for a simulated time series based on a quadratic increase with added noise error, which is normal distributed with a mean of zero and a standard deviation of one. The total increase at the end of the time series is equal to 50 standard deviations. The filled bar indicates the best AIC/BIC model.

It can be observed that both information criteria suggest a model based on a linear trend, what is consistent with the results of EnvCpt for the pCO_2 time series. Moreover, the trend models describe the time series using piecewise linear periods resulting in a high number of changepoints. This explains why in a lot of regions the number of detected changepoints in the pCO_2 time series is relatively high, as can be seen in figure 5(d).

Another thing that can be observed is that with a bigger difference between the start and the end point of the quadratic increase, the models considering second-order auto-correlation (AR2) are not detecting changepoints anymore and the large increase is fully described with the autocorrelation. However, the support of these AR models for the time series is relatively bad and EnvCpt will still suggest a trend model with changepoints.

4.2.3 Higher-order Autocorrelation

In a last simulation the behavior of EnvCpt is studied when the data is based on higher-order autocorrelation. For that, data with different orders of autocorrelation has been simulated and again, normal distributed noise error with a mean of zero and a standard deviation of one was added. Since no additional behavior is added to the time series, the best model suggested by EnvCpt should be the mean model with respective order autocorrelation.

In order to see if EnvCpt performs well, the best AIC/BIC model was determined and counted for 1000 repetitions. The results are presented in bar-plots showing the distribution of the suggested models in figure 20. This analysis was done with autocorrelation of first and second-order, which are also taken into account by Envcpt, and third-order, which is not included anymore.

First-order autocorrelation:

Using AIC as the information criterion shown in plot 20(a), the trend model with AR2 is the best model in about 850 cases. Since the autocorrelation leads to a dependence between the time steps, a longer period of an increase or a decrease can occur. EnvCpt often tends to interpret such behavior as a linear trend, although no underlying trend was added. Using the BIC, the mean model with AR1 has been the best model in most of the cases. The difference again lies in the penalty term of the information criterion. Since the trend models are penalized more because of the larger amount of free adjustable parameters, the resulting best model is a mean model. As mentioned before, the large differences between AIC and BIC plots again show that the log-likelihood of some models are in the same range, meaning that they have similar support.

Second-order autocorrelation:

By looking at the result of the AIC in plot 20(c) it can be observed that there are only small differences to the AR1 plot. However, since the time series is based on a AR2 process, AIC tends to pick the correct order of autocorrelation this time. Yet, it still infers a trend that is not present in the toy data. With the BIC, the best models are often models considering AR2, and unlike AIC, BIC suggest the correct mean AR2 model in some cases.
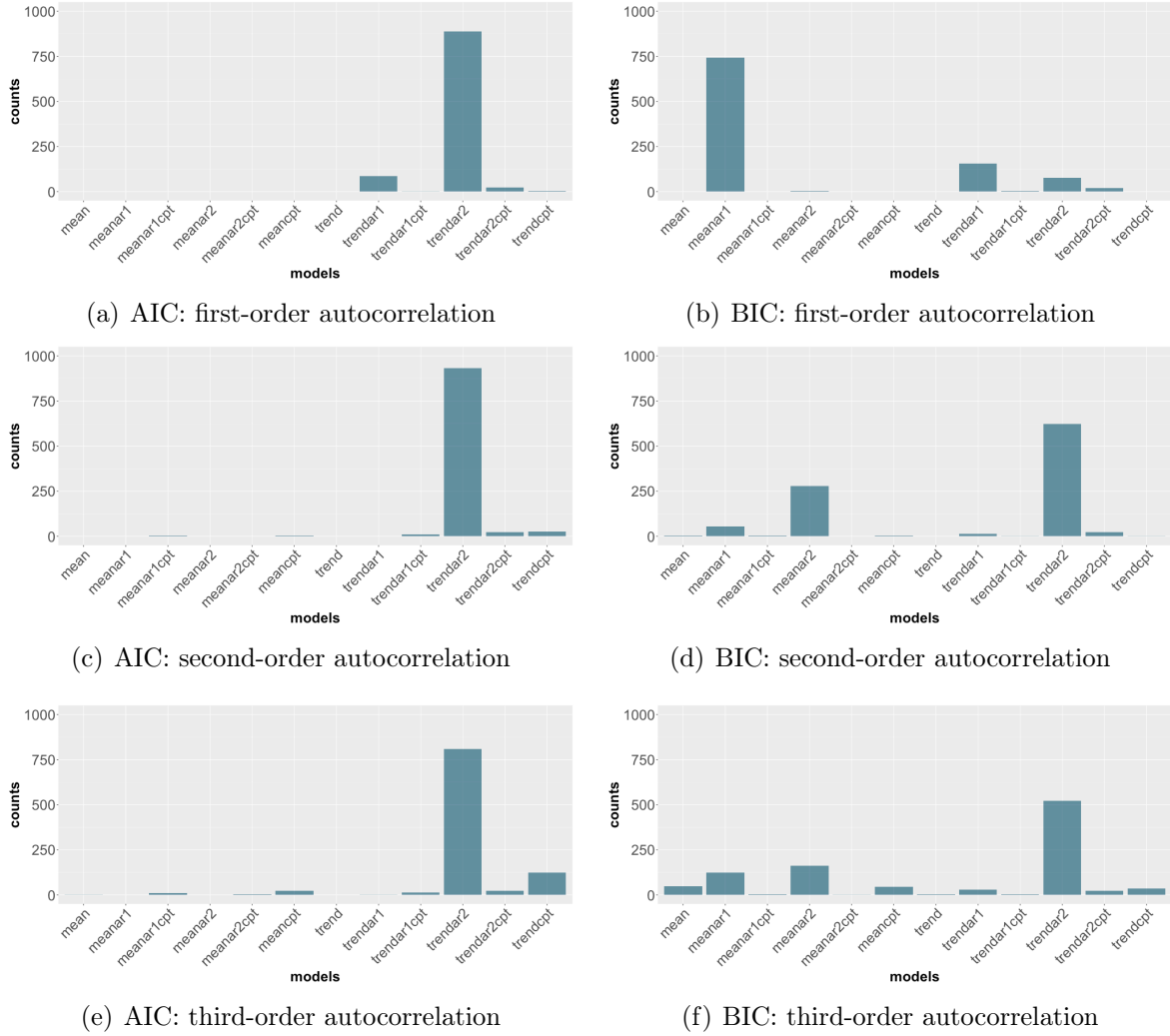
(a) AIC: first-order autocorrelation

(b) BIC: first-order autocorrelation

(c) AIC: second-order autocorrelation

(d) BIC: second-order autocorrelation

(e) AIC: third-order autocorrelation

(f) BIC: third-order autocorrelation

Figure 20: Distribution of the best model detected with both AIC and BIC for time series based on autocorrelation of order one, two and three with added normal distributed noise error with a mean of zero and a standard deviation of one. In all plots the best model has been counted for 1000 repetitions.

Third-order autocorrelation:

With third order autocorrelation, the best model is the trend model considering AR2 in most cases. In contrast to the first and second orders, the suggestion of non-autocorrelation models with changepoints is higher.

4.2.4 Limitations of EnvCpt

The performance of EnvCpt with simulated toy data shows that there are some clear limitations in the detection of changepoints. First of all, EnvCpt is only considering a selection of statistical models. Changes in the time series following increases higher-than-linear order or higher order autocorrelation can lead to a misinterpretation of the data. The study of the quadratic increase showed that such behavior results in the detection of a lot of changepoints, although the underlying data follows a continuous curve. The

presence of autocorrelation has been detected correctly in most cases, even at higher order than considered in EnvCpt. Nevertheless, the use of EnvCpt for the detection of changepoints is limited to time series, which behavior is included in EnvCpt. Therefore it is important to know the general changes in the time series before interpreting the results of the changepoint analysis with EnvCpt.

Using a weight condition to determine how valid the detection of a changepoint is can lead to the loss of informations. As explained in section 4.2.1, if several models have similar AIC/BIC, the weight is distributed on these models. This problem can be solved by adding up the weight of changepoints with similar position.

5   Conclusions

Analyzing the first results, we noticed that EnvCpt detected changepoints in almost every grid cell. In order to handle this large amount of data, two different filtering methods have been introduced. In both methods a condition on the weight of either the changepoint models or the changepoints themself are defined. On one hand, the weight condition helps to reduce the amount of changepoints and on the other hand, only changepoint models or changepoints with a certain support for the time series are considered.

With the results based on the default parameter choice, several regions with a high rate of changepoints could be identified. Often, the changes occur in the RCP8.5 period as for example in high-latitude regions, where the decline of the sea ice concentration could only be observed starting in the middle of the 21$^{\text{st}}$ century. However, also bimodal switches in sea ice concentration as well as in the anticorrelated net primary production of phytoplankton throughout the historical period could be observed, which have also been found by Drijfhout et al. (2015). Moreover, Drijfhout et al. (2015) found abrupt shifts leading to rapid changes to a new state. While this is also simulated in our data (as for example in the Barents Sea), the used plots gave no clear indication for these kind of abrupt changes and therefore, they would not have been detected in the analysis with EnvCpt.

Besides the different methods, the definitions of abrupt changes [Drijfhout et al., 2015] and changepoints [Beaulieu C., Killick R., 2018] are not the same. While an abrupt change is related to a time scale and is defined to be occurring faster than the time scale of the external forcing [Drijfhout et al., 2015], the changepoints detected in EnvCpt are describing changes in the statistical properties of a time series. EnvCpt, or more precisely the Optimal Partioning algorithm implemented in EnvCpt, relates the changes in time series to the periods before and after the change. If the behavior of the change period differs sufficiently enough (decided by the cost function) from the behavior before or after, EnvCpt decides that the data is better described by changepoint models. In this method, the changepoint is not related to an absolute value of the change over a certain time period (in contrast to Drijfhout et al. (2015)), which leads to the ability of EnvCpt to detect different types changepoints (mean or trend). Therefore, the concept used in EnvCpt is more general than the approach by Drijfhout et al. (2015).

However, since the changepoints are not related to an absolute value of change, the following needs to be noted. If a variable changes the same way over an entire time series, EnvCpt will not detect changepoints, provided the type of change is included in EnvCpt. EnvCpt only detects changes in the statistical properties and if for example a variable follows exactly a linear trend, no changepoints would result. So, if for example the type 1 plot does not indicate a high rate of changepoints, it can not be concluded that the variable does not underlie any changes. This shows again that the results are strongly influenced by the choice of the segment of the analyzed time series, since changes are only detected when the statistical properties are changing.

Moreover, it should be mentioned that there is no disadvantage from the fact that EnvCpt does not combine several models for the fit on one time series. This can be seen by the following considerations. If a time series consists of two mean periods interrupted by a linear-trend period, EnvCpt will likely suggest the trend model, whereby the slope in

the mean periods is set to zero. The same can be applied on a non-autocorrelation time series, which is divided by a period based on autocorrelation. Since the autocorrelation coefficients in the respective models can be estimated as zero, EnvCpt will probably describe the time period by a model of the respective order of the autocorrelation period. This also shows that if EnvCpt would neglect parameters that are estimated close to zero, considering mean and non-autocorrelation models would be redundant.

The two filter methods as well as the three plot types based on them processed the EnvCpt output in such a way that some conclusions could be drawn about changes in the variables. However, in order to get a deeper understanding of the processes leading to abrupt changes, further investigations are needed. Especially the plots showing the changepoints that occurred in a single year are delivering relevant informations about the correlation between the changes in the analyzed variables.

Moreover, it could be seen that changing the input parameters can have a large influence on the results. Especially changing the information criterion along with a defined weight condition can have major impact on the detected changepoints. While by using the BIC clearer patterns are resulting due to the larger penalty term, in other regions no changepoints at all are detected, much different from the results obtained with AIC. Since the AIC is less penalizing and therefore, the effect of information loss resulting from the combination of the higher penalty term in the BIC and the weight condition (as observed in the Weddell Sea, see section 3.3.1) is less likely, the AIC is suggested as the default information criterion. Nevertheless, the BIC can be useful in regions where the AIC shows no clear pattern.

Using toy data, some limitations of EnvCpt, as for example the falsely detected changepoints when a time series follows a higher-than-linear-order behavior, could be pointed out. However, another important aspect needs to be studied. As mentioned in the method section (section 2.2), the residuals obtained by fitting a model to a time series should be normal distributed and independent. Only after analyzing the residuals it can be determined whether a model suggested by EnvCpt actually delivers a substantial support for the time series or not. Nevertheless, since models considering a constant mean or a linear trend are often sufficient to describe variables representing ocean or atmosphere data (there are exceptions as for example $p\text{CO}_2$), it can be expected that these requirements are met in the most cases in this analysis.

6   Acknowledgments

References

[H. Aikake, 1974] Aikake H., *A new look at the statistical model identification*, in IEEE Transactions on Automatic Control, vol. 19, no. 6, pp. 716-723, December 1974, doi: `https://doi.org/10.1109/TAC.1974.1100705`

[Beaulieu C., Killick R., 2018] Beaulieu C., Killick R., *Distinguishing trends and shifts from memory in climate data*, in Journal of Climate, vol. 31, pp. 9519–9543, December 2018, doi: `https://doi.org/10.1175/JCLI-D-17-0863.1`

[Burger et al., 2020] Burger F. A., Frölicher T. L., John J. G.: *Increase in ocean acidity variability and extremes under increasing atmospheric* CO_2, in Biogeosciences Discussions, vol. 2020, pp. 1-32, January 2020, doi: `https://doi.org/10.5194/bg-2020-22`

[Drijfhout et al., 2015] Drijfhout S. et al., *Catalogue of abrupt shift in Intergovernmental Panel on Climate Change climate models*, in Catalogue of abrupt shifts in Intergovernmental Panel on Climate Change climate models, vol. 112, no. 43, pp. E5777-E5786, September 2015, doi: `https://doi.org/10.1073/pnas.1511451112`

[Drijfhout et al., 2012] Drijfhout S., Van Oldenborgh G. J., Cimatoribus A., *Is a Decline of AMOC Causing the Warming Hole above the North Atlantic in Observed and Modeled Warming Patterns?*, in Journal of Climate, vol. 25, no. 24, pp. 8373-8379, December 2012, doi: `https://doi.org/10.1175/JCLI-D-12-00490.1`

[Dunne et al., 2012] Dunne J. P. et al., *GFDL's ESM2 Global Coupled Climate–Carbon Earth System Models. Part I: Physical Formulation and Baseline Simulation Characteristics*, in Journal of Climate, vol. 25, no. 19, pp. 6646-6665, April 2012, doi: `https://doi.org/10.1175/JCLI-D-11-00560.1`

[ESM GFDL, accessed 16.05.2020] *Earth System Model - Geophysical Fluid Dynamics Laboratory*, accessed 16.05.2020, url: `https://www.gfdl.noaa.gov/earth-system-model/`

[Haynes et al., 2015] Haynes K., Eckley I. A., Fearnhead P., *Computationally Efficient Changepoint Detection for a Range of Penalties*, in Journal of Computational and Graphical Statistics, vol. 26, no. 1, pp. 134-143, December 2015, doi: `https://doi.org/10.1080/10618600.2015.1116445`

[Killick et al., 2012] Killick R., Fearnhead P., Eckley I. A., *Optimal Detection of Changepoints With a Linear Computational Cost*, in Journal of the American Statistical Association, vol. 107, no. 500, pp. 1590-1598, October 2012 , doi: `https://doi.org/10.1080/01621459.2012.737745`

[Taylor et al., 2012] Taylor K., Ronald, S., Meehl G., *An overview of CMIP5 and the Experiment Design*, in Bulletin of the American Meteorological Society, vol. 93, pp. 485-498, November 2012, doi: `https://doi.org/10.1175/BAMS-D-11-00094.1`

[Zhang N., Siegmund D., 2007] Zhang N. R., Siegmund D. O., *A Modified Bayes Information Criterion with Applications to the Analysis of Comparative Genomic Hybridization Data*, in Biometrics, vol. 63, pp. 22-32, April 2007, `https://10.1111/j.1541-0420.2006.00662.x`

A   R scripts

Listing 1: Evaluation of `envcpt` for each grid cell

```r
for(x in 1:360) {
  for(y in 1:200) {
  data_local <- data[x,y,] #get the time series of the grid cell
  #check for missing data
  if (anyNA(data_local)==FALSE && max(data_local)!=0) {
    data_local_envcpt <- envcpt(data_local) #evaluate envcpt
    #get all AIC informations
    aic <- AIC(data_local_envcpt)
    delta_aic <- aic-min(aic,na.rm=T)
    min_aic <- which.min(aic)
    aic_weight <- AICweights(data_local_envcpt)
    #collecting informations of the best AIC model (same for the BIC)
    best_name_aic <- names(min_aic)
    best_nr_aic <- unname(min_aic)
    best_aic_weight <- unname(aic_weight[best_nr_aic])
    best_cpts_aic <- NA
    if (best_nr_aic %in% cpts_mod) {
      best_cpts_aic <- data_local_envcpt[[best_name_aic]]@cpts
    }
    best_aic_model <- list(model_nr=best_nr_aic, model_name=best_name_aic,
                  aic_weight=best_aic_weight, cpts=best_cpts_aic)
    #collecting informations of all the models
    info <- NULL
    for (i in 1:12) {
    if(i %in% cpts_mod) {
      change_pnts <- data_local_envcpt[[model_names[i]]]@cpts
      nr_of_cpts <- length(change_pnts)
    } else {
      nr_of_cpts <- 0
    }
    info <- rbind(info, data.frame(i, model_names[i], aic[i], bic[i],delta_aic[i],
                  delta_bic[i], aic_weight[i], bic_weight[i],nr_of_cpts))
    }
  #collecting all the changepoint data
  ch_pnts_data <- NULL
  for(i in cpts_mod) {
    ch_pnts <- data_local_envcpt[[model_names[i]]]@cpts #get the changepoints
    nr_cpts <- length(ch_pnts)
    if(nr_cpts != 0) {
      for (j in 1:nr_cpts) {
        ch_pnts_data <- rbind(ch_pnts_data, data.frame(i, model_names[i],
                                                ch_pnts[j], aic_weight[i],
            bic_weight[i]))
      }
    }
  }
  #storing all informations in a list
  if(!is.null(ch_pnts_data)) {
    ch_pnts_data <- ch_pnts_data[order(ch_pnts_data$position),]
    nr_cpts <- length(ch_pnts_data$position)
    longitude <- lon[x,y]
    latitude <- lat[x,y]
    #filling list with a list for every grid cell with changepoints
    result[[cnt]] <- list(grid_cell= c(x,y),coordinates=c(longitude,latitude),
                        infos=info,best_aic_model=best_aic_model,
                        best_bic_model=best_bic_model, change_points=ch_pnts_data,
                        number_of_cpts=nr_cpts)
    }
    }
  }
}
```

Listing 2: Filtering the output of EnvCpt by the best AIC/BIC model (method I)

```
1   for(i in 1:length(result)) {
2     #delete changepoints which are not in the defined time interval (same for the BIC)
3     if(!anyNA(result_fil[[i]][["best_aic_model"]]$cpts)) {
4         del_cpts <- NULL
5         del_cpts_cnt <- 1
6         for(j in 1:length(result_fil[[i]][["best_aic_model"]]$cpts)) {
7             if(result_fil[[i]][["best_aic_model"]]$cpts[j] %nin% search_start:search_end) {
8               del_cpts[del_cpts_cnt] <- j
9               del_cpts_cnt <- del_cpts_cnt+1
10            }
11            if(!is.null(del_cpts)) {
12              result_fil[[i]][["best_aic_model"]]$cpts <- result_fil[[i]][["best_aic_model"
                                                ]]$cpts[-del_cpts]
13            }
14        }
15      }
16    #if all changepoints were deleted, cpts is changed to NA
17    if(length(result_fil[[i]][["best_aic_model"]]$cpts)==0) {
18      result_fil[[i]][["best_aic_model"]]$cpts <- NA
19    }
20    #set cpts to NA if the best AIC weight is not large enough (same for the BIC)
21    if(!is.na(result_fil[[i]][["best_aic_model"]][["best_aic_weight"]])) {
22      if(result_fil[[i]][["best_aic_model"]][["aic_weight"]] < min_weight) {
23        result_fil[[i]][["best_aic_weight"]]$cpts <- NA
24      }
25    } else {
26      result_fil[[i]][["best_aic_model"]]$cpts <- NA
27    }
28    #delete grid cells, where the number of changepoints is NA
29    if(anyNA(result_fil[[i]][["best_aic_model"]]$cpts)) {
30      del_grid[del_grid_cnt] <- i
31      del_grid_cnt <- del_grid_cnt+1
32    }
33  }
34  #update the result_fil list
35  if (!is.null(del_grid)) {
36    result_fil <- result_fil[-del_grid]
37  }
```

Listing 3: Filtering the output of EnvCpt and cluster changepoints (method II)

```
1   for(i in 1:length(result))
2     #cluster changepoints starting with the best AIC weight model (same for the BIC)
3     if(result_fil[[i]]$number_of_cpts != 0) {
4       del_cpts <- NULL
5       del_cpts_cnt <- 1
6       #sort changepoint list after AIC weight
7       result_fil[[i]]$change_points <- result_fil[[i]]$change_points[order(result_fil[[i]]
                                          $change_points[[aic_weight]], decreasing = TRUE),]
8       first_cpt_model <- result_fil[[i]]$change_points[1,1]
9       #get list with changepoints of the best AIC model
10      for (j in 1:length(result_fil[[i]]$change_points$position)) {
11        if(result_fil[[i]]$change_points$model_nr[j] != first_cpt_model) {
12          del_cpts[del_cpts_cnt] <- j
13          del_cpts_cnt <- del_cpts_cnt+1
14        }
15      }
16      if (!is.null(del_cpts)) {
17        added_cpts_data <- result_fil[[i]]$change_points[-del_cpts,]
18      } else {
19        added_cpts_data <- result_fil[[i]]$change_points
20      }
21      #go through the rest of the list, which is already sorted after the AIC weight
22      for(j in 1:length(result_fil[[i]]$change_points$position)) {
23        if(result_fil[[i]]$change_points$model_nr[j] != first_cpt_model) {
24          added <- FALSE
25          for(k in 1:length(added_cpts_data$position)) {
26            if (result_fil[[i]]$change_points$position[j] %in% (added_cpts_data
                  $position[k]-int_cluster):(added_cpts_data$position[k]+int_cluster)) {
27              added_cpts_data$aic_weight[k] <- added_cpts_data$aic_weight[k] +
                                                result_fil[[i]]$change_points$aic_weight[j]
```

```
28              added <- TRUE
29              break
30            }
31          }
32          if (added == FALSE) {
33            added_cpts_data <- rbind(added_cpts_data, result_fil[[i]]$change_points[j,])
34          }
35        }
36      }
37      #delete changepoints which don't fulfill the total AIC condition (same for the BIC)
38      if(result_fil[[i]]$number_of_cpts!=0) {
39        del_cpts <- NULL
40        del_cpts_cnt <- 1
41        for(j in 1:result_fil[[i]]$number_of_cpts) {
42          if(is.na(result_fil[[i]]$change_points[[total_aic_weight]][j]) ||
43             result_fil[[i]]$change_points[[paste(total_aic_weight]][j] < min_weight) {
44            del_cpts[del_cpts_cnt] <- j
45            del_cpts_cnt <- del_cpts_cnt+1
46          }
47        }
48        if(!is.null(del_cpts)) {
49          result_fil[[i]]$change_points <- result_fil[[i]]$change_points[-del_cpts,]
50          result_fil[[i]]$number_of_cpts <- result_fil[[i]]$number_of_cpts-length(del_cpts)
51        }
52      }
53      #delete changepoints which are not in the defined time interval
54      if(result_fil[[i]]$number_of_cpts!=0) {
55        del_cpts <- NULL
56        del_cpts_cnt <- 1
57        for(j in 1:result_fil[[i]]$number_of_cpts) {
58          if(result_fil[[i]]$change_points$position[j] %nin% search_start:search_end) {
59            del_cpts[del_cpts_cnt] <- j
60            del_cpts_cnt <- del_cpts_cnt+1
61          }
62        }
63        if(!is.null(del_cpts)) {
64          result_fil[[i]]$change_points <- result_fil[[i]]$change_points[-del_cpts,]
65          result_fil[[i]]$number_of_cpts <- result_fil[[i]]$number_of_cpts-length(del_cpts)
66        }
67      }
68 }
```

Listing 4: Simulating toy data with polynomial behavior and noise error

```
1  cpts <- TRUE
2  pos <- c(120)
3  ntot <- 240
4  std <- 1
5  #polynomial function (a+bx+cx^2) for every interval
6  a <- c(0,4*std)
7  b <- c(0,0)
8  c <- c(0,0)
9  #create (piecewise) polynomial function
10 if(isTRUE(cpts)) {
11   pos <- c(0,pos,ntot)
12   y <- NULL
13   for(i in 1:(length(pos)-1)){
14     x <- (pos[i]+1):pos[i+1]
15     y_temp <- a[i]+b[i]*x+c[i]*x^2
16     y <- c(y, y_temp)
17   }
18 } else {
19   x <- 1:ntot
20   y <- a + b*x + c*x^2
21 }
22 noise_error <- arima.sim(list(order = c(0,0,0)), sd=std, n=ntot)
23 toy_data <- y + noise_error
24 data_envcpt <- envcpt(toy_data)
25 layout(matrix(c(1,2,3), 1, 3, byrow = TRUE), widths=c(3,1,1), heights=c(1))
26 plot(data_envcpt, type=c("fit"), cex.axis=1.7)
27 plot(data_envcpt, type=c("aic"), cex.axis=1.7)
28 plot(data_envcpt, type=c("bic"), cex.axis=1.7)
```

Listing 5: Determining the distribution of the best AIC/BIC model

```
1   ntot <- 240
2   std <- 1
3   rep <- 1000
4   ord <- 3
5   #create dataframe to store the counters
6   model_name <- c("mean","meancpt","meanar1","meanar2","meanar1cpt","meanar2cpt",
7                   "trend","trendcpt","trendar1","trendar2","trendar1cpt","trendar2cpt")
8   counts <- data.frame(model_name, rep(0, each=12), rep(0, each=12))
9   #define the AR coefficients
10  ar_cof <- rep(0.5/ord, each=ord)
11  for(i in 1:rep) {
12    noise_error <- arima.sim(list(order = c(0,0,0)), sd=std, n=ntot)
13      data <- arima.sim(list(order = c(ord,0,0), ar=ar_cof), sd=std, n=ntot) + noise_error
14      data_envcpt <- envcpt(data)
15      aic <- AIC(data_envcpt)
16      bic <- BIC(data_envcpt)
17      min_aic <- unname(which.min(aic))
18      min_bic <- unname(which.min(bic))
19      counts$aic_counts[min_aic] <- counts$aic_counts[min_aic] +1
20      counts$bic_counts[min_bic] <- counts$bic_counts[min_bic] +1
21  }
```

B Additional plots

B.1 Robustsness of the Results

Type 1: Total number of detected changepoints

(a) sea surface temperature

(b) sea ice concentration

(c) net primary production
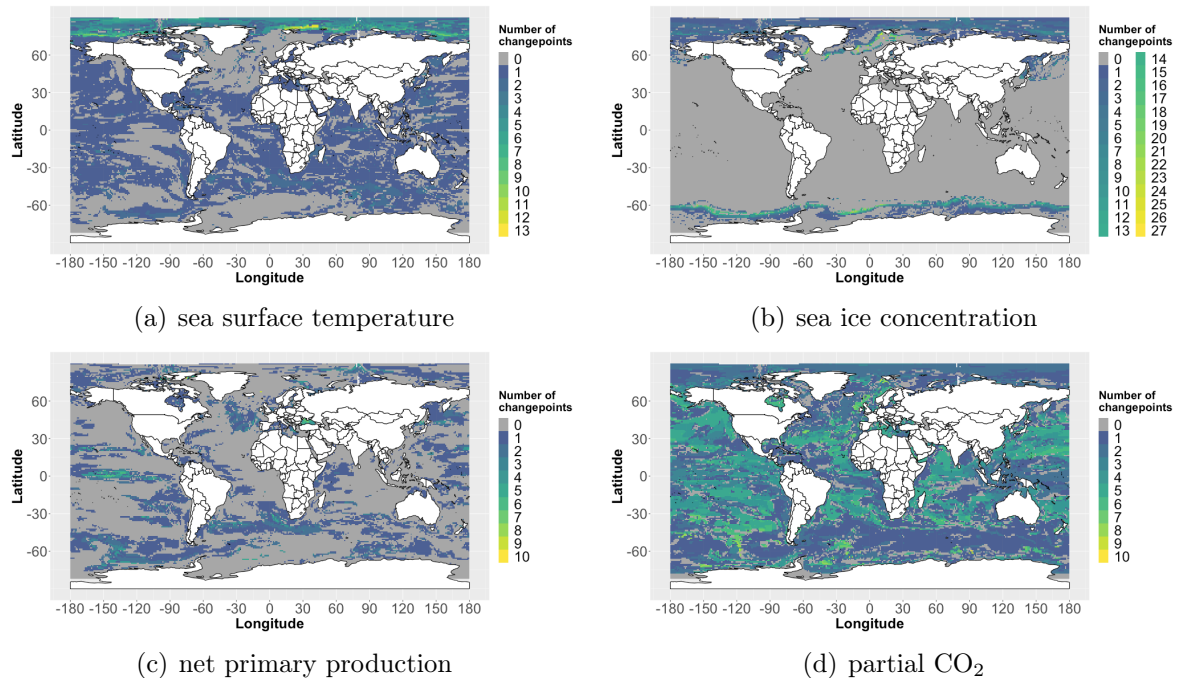
(d) partial CO_2

Figure 21: The number of changepoints for the four analyzed variables based on the default parameters with BIC instead of AIC and the time series containing the historical as well as the RCP8.5 period.

Type 2: Nature of the changepoints

(a) sea surface temperature

(b) sea ice concentration

(c) net primary production
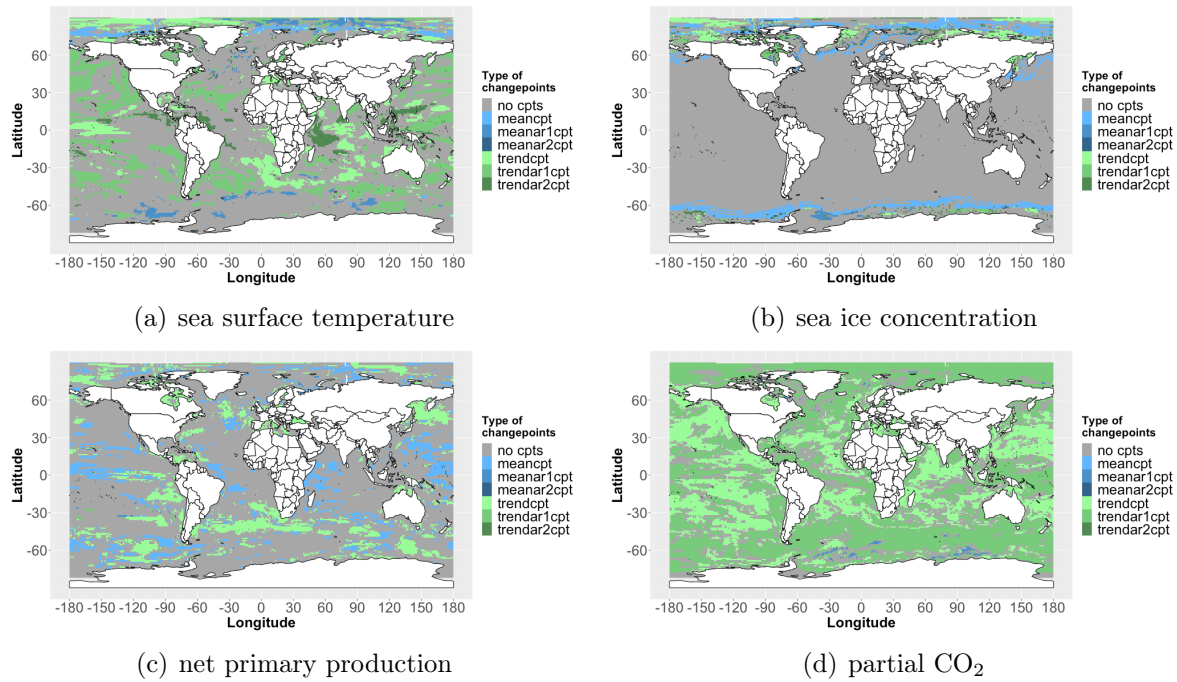

(d) partial CO_2

Figure 22: The nature of changepoints for the four analyzed variables based on the default parameters with BIC instead of AIC and the time series containing the historical as well as the RCP8.5 period.

Type 3: Year of the first changepoint

(a) sea surface temperature

(b) sea ice concentration

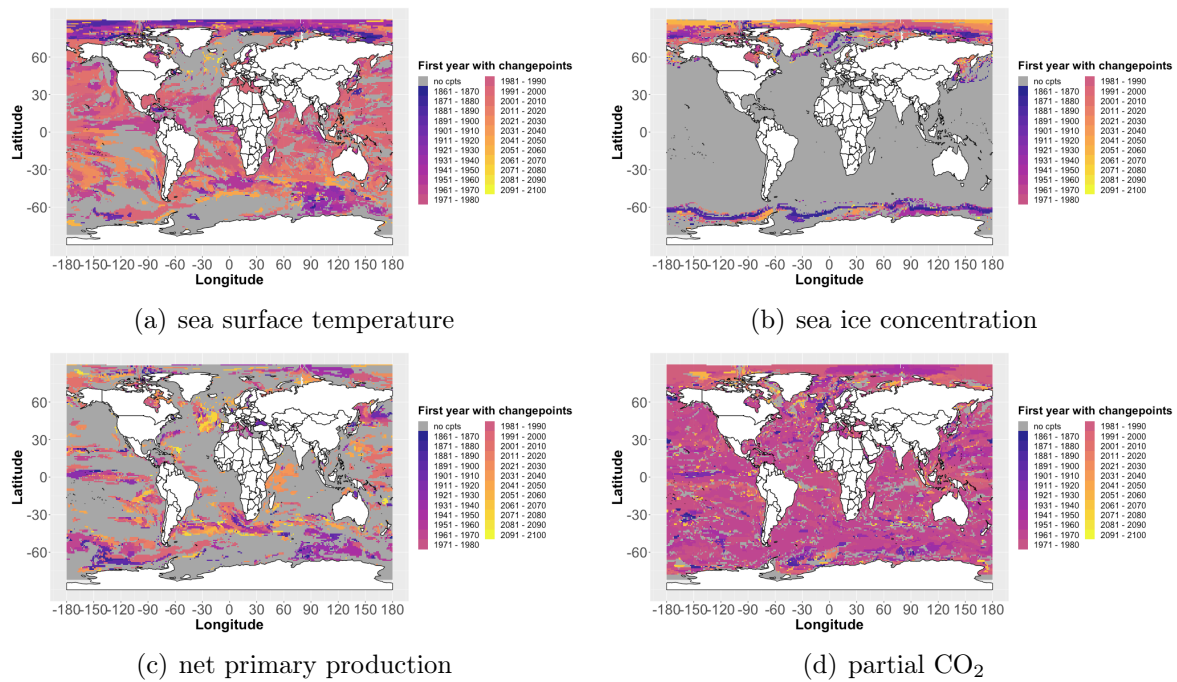(c) net primary production

(d) partial CO_2

Figure 23: Year of first changepoint for the four analyzed variables based on the default parameters with BIC instead of AIC and the time series containing the historical as well as the RCP8.5 period.