



# A regional $p\text{CO}_2$ climatology of the Baltic Sea from in situ $p\text{CO}_2$ observations and a model-based extrapolation approach

Henry C. Bittig, Erik Jacobs, Thomas Neumann, and Gregor Rehder

Leibniz Institute for Baltic Sea Research Warnemünde (IOW), Rostock-Warnemünde, Germany

**Correspondence:** Henry C. Bittig (henry.bittig@io-warnemuende.de)

Received: 6 July 2023 – Discussion started: 25 July 2023

Revised: 4 December 2023 – Accepted: 10 December 2023 – Published: 6 February 2024

**Abstract.** Ocean surface  $p\text{CO}_2$  estimates are of great interest for the calculation of air–sea  $\text{CO}_2$  fluxes, oceanic uptake of anthropogenic  $\text{CO}_2$ , and eventually the Global Carbon Budget. They are accessible from direct observations, which are discrete in space and time and thus always sparse, or from biogeochemical models, which only approximate reality. Here, a combined method for the extrapolation of  $p\text{CO}_2$  observations is presented that uses (1) model-based patterns of variability from an empirical orthogonal function (EOF) analysis of variability with (2) observational data to constrain EOF pattern amplitudes in (3) an ensemble approach, which locally adjusts the spatial scale of the mapping to the density of the observations. Thus, data-constrained, gap- and discontinuity-free mapped fields including local error estimates are obtained without the need for or dependence on ancillary data (e.g. satellite sea surface temperature maps). This extrapolation approach is generic in that it can be applied to any oceanic or coastal region covered by a suitable model and observations. It is used here to establish a regional  $p\text{CO}_2$  climatology of the Baltic Sea (Bittig et al., 2023: <https://doi.org/10.1594/PANGAEA.961119>), largely based on ICOS-DE ship of opportunity (SOOP) *Finnmaid* surface  $p\text{CO}_2$  observations between Lübeck-Travemünde (Germany) and Helsinki (Finland). The climatology can serve as improved input for atmosphere–ocean  $\text{CO}_2$  flux estimation in this coastal environment.

## 1 Introduction

The ocean plays a major role in the global carbon cycle and has a controlling function on the atmospheric  $\text{CO}_2$  content on longer timescales (DeVries, 2022). Since the rise of atmospheric  $\text{CO}_2$  concentrations since the beginning of the pre-industrial period (year 1750), the ocean has taken up  $\sim 25\%$  of the  $\text{CO}_2$  released from human activities (Friedlingstein et al., 2022), with the annual uptake mainly related to the increase in the air–sea  $p\text{CO}_2$  imbalance. The role of coastal and continental shelf waters is more complex. Apart from atmospheric  $\text{CO}_2$  levels, changes in nutrient loads and organic matter supply from land, changes in weathering in the drainage basin, and even changes in the functioning and composition of biological key players on various levels can lead to changes in the inorganic carbon system and thus the source–sink function of coastal seas (e.g. Laruelle

et al., 2018; Müller et al., 2016; Carstensen and Duarte, 2019; Kuliński et al., 2022). Moreover, coastal seas provide an important conduit of land-derived carbon into the open ocean's interior (e.g. Thomas et al., 2004).

For the Baltic Sea, several attempts have been made to quantify the net  $\text{CO}_2$  air–sea balance in the form of a  $p\text{CO}_2$  climatology as well as to derive trends in surface water  $p\text{CO}_2$ , with partly contradictory results (e.g. Omstedt et al., 2009; Parard et al., 2016, 2017; Becker et al., 2021; Neumann et al., 2022; Wesslander et al., 2010). Most of the approaches either used the output from biogeochemical models or tried to create  $p\text{CO}_2$  fields from mapped proxy data and observational data using extrapolation approaches. Seasonal mapping of  $p\text{CO}_2$  is particularly challenging for the Baltic Sea due to its high regional and temporal variability, a salinity gradient affecting e.g.  $\text{CO}_2$  equilibria, and a large seasonal amplitude caused by high net productivity in spring and summer and

entrainment of waters enriched in remineralization products due to mixed-layer deepening in autumn and winter (Schneider and Müller, 2018).

Climatologies of  $p\text{CO}_2$  on a global or ocean-wide scale are an important tool for the quantification of the oceanic  $\text{CO}_2$  sink in the framework of e.g. the Global Carbon Budget (Landschützer et al., 2013; Friedlingstein et al., 2022). Tailored regional analysis can help to gain insight into changes in the source–sink behaviour of distinct regions, as recently demonstrated for the northern European shelf, including the Baltic Sea (Becker et al., 2021).

A robust climatology and trend for the Baltic Sea has, apart from refining the estimate of the net  $\text{CO}_2$  flux in relation to the Global Carbon Budget, several implications. As the entire Baltic Sea area belongs to the territorial waters or exclusive economic zones (EEZs) of one of the pan-Baltic nations, the air–sea  $\text{CO}_2$  fluxes might be of importance for current and future carbon management and accounting schemes in the framework of national emission reduction targets (Luisetti et al., 2020). A  $p\text{CO}_2$  climatology could also serve as a baseline for potential negative emissions applications, including blue carbon or coastal alkalinity enhancement (GESAMP, 2019). Knowledge of monthly  $p\text{CO}_2$  fields and their variability might also help to identify and quantify the impact of perturbations and extreme events, e.g. heat waves (Humborg et al., 2019).

In this work, we build a foundation for such applications. We first present a novel extrapolation approach, followed by construction of a Baltic Sea  $p\text{CO}_2$  climatology. The extrapolation approach is then evaluated and put into context with existing mapping methods. Notable features of the seasonal  $p\text{CO}_2$  climatology are discussed and special attention given to the regional long-term  $p\text{CO}_2$  trend before we conclude our work.

## 2 Methods

### 2.1 Extrapolation approach

For mapping from scarce observational data to spatially filled maps of the Baltic Sea, we use an ensemble of truncated empirical orthogonal function (EOF) reconstructions. For a more detailed description than the brief summary below, please consult Appendix A.

EOF decomposition and singular value decomposition (SVD) have been used widely in atmospheric and ocean science (e.g. Lorenz, 1956; Weare et al., 1976; Weare and Newell, 1977; Hannachi et al., 2007). They can be used to efficiently reduce the dimensionality of the original dataset (Lorenz, 1956; Davis, 1976; Preisendorfer, 1988; Monahan et al., 2009; Jolliffe and Cadima, 2016). Here, we use the EOF decomposition of a model dataset  $\mathbf{X}$  of  $p\text{CO}_2$  in the Baltic Sea to obtain spatial EOF patterns  $e_i$  of  $p\text{CO}_2$  variability.

In a second step, observational  $p\text{CO}_2$  data  $\mathbf{Y}$  are used in conjunction with these spatial patterns of variability in a truncated EOF reconstruction to constrain the EOF spatial patterns' amplitudes (Kaplan et al., 2000; Preisendorfer, 1988) and thus to extrapolate from scarce observational data to the full domain (Sect. A3). This step represents an optimization of a cost function  $Q$  (Eq. A19), which in our case also takes the observational data uncertainty into account to avoid overfitting (Sect. A4). For a given EOF reconstruction with  $l$  spatial modes, we thus obtain both an extrapolated field as well as an estimate of the mapping uncertainty  $\sigma_P$ , representing the uncertainty in the extrapolation with the given data constraints. In addition, we derive a representational error estimate  $\sigma_{P'}$  for the variance not resolved by the given EOF reconstruction due to truncation and use of just  $l$  modes (Sect. A4).

However, the choice of how many EOF modes to use for a given truncated EOF reconstruction, or at which level to truncate, is an arbitrary one often inspired by a certain threshold of the total variance explained (e.g. Kaplan et al., 2000).

In a third step, we therefore use an ensemble approach to circumvent this problem: instead of a single EOF reconstruction with a fixed number of modes  $l$ , we use a series of EOF reconstructions from just one mode up to a maximum number of  $l_{\text{max}}$  modes (Sect. A5). This series is then combined by a weighted ensemble at each location with the individual EOF reconstruction's mapping uncertainties as weights (Eqs. A25 and A26; Sect. A6).

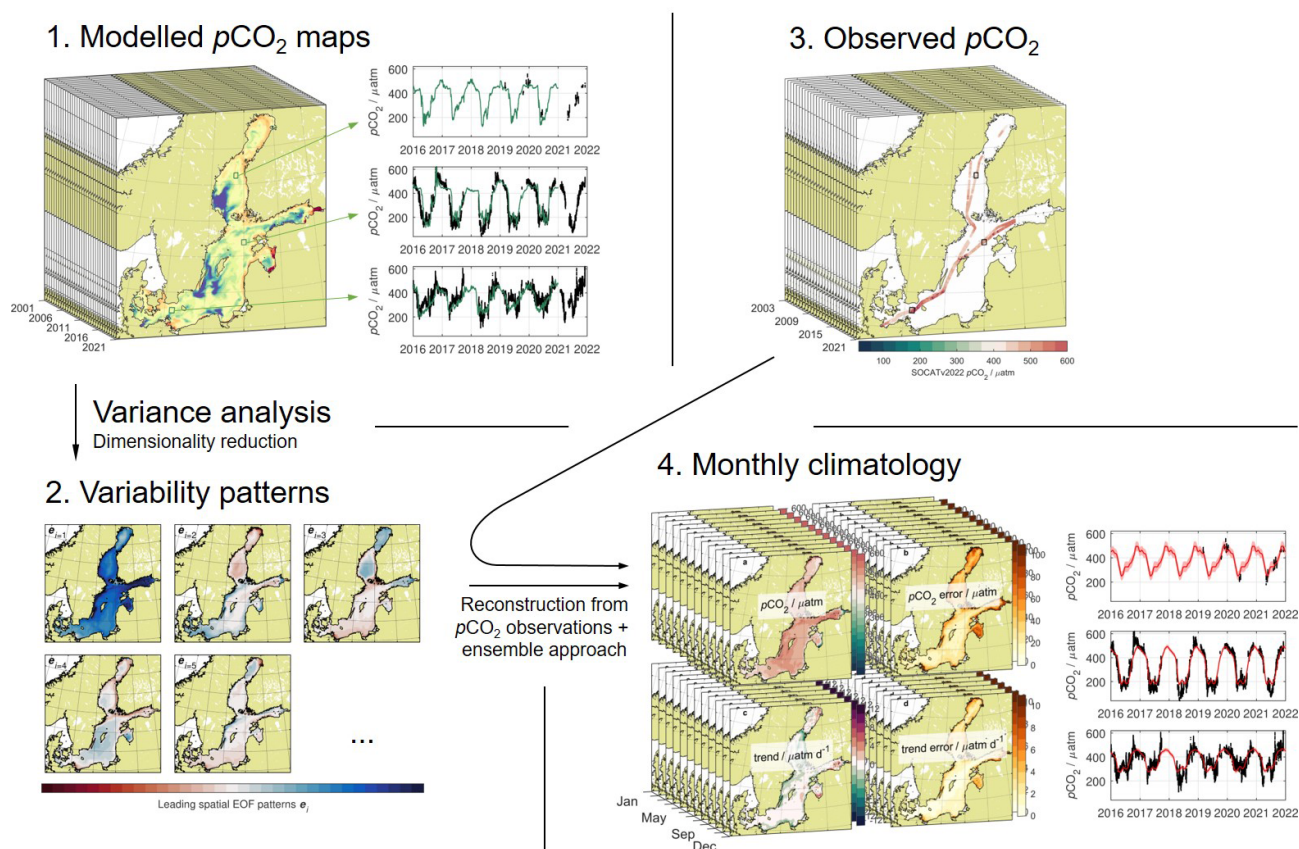
Note that the weights are spatially resolved and depend on the available data constraints, i.e. the ensemble weights provide for locality, which includes adaptation of the mapping's spatial scales to the data constraint density, thus providing for a more robust extrapolation and more realistic uncertainty estimates than with a fixed number  $l$  of EOF modes.

We thus obtain an ensemble mean  $p\text{CO}_2$  value  $x_{\text{reconstr}}$ , the average number of modes  $\bar{l}$  used in the ensemble of reconstructions at a given location, and an uncertainty estimate  $\sigma_{\text{reconstr}}$  (Eq. A28), which consists of the sample variance about the ensemble mean value, the ensemble-averaged mapping uncertainty  $\sigma_P$  (from each truncated EOF reconstruction with  $l$  from  $1 \dots l_{\text{max}}$ ), and the ensemble-averaged representational error  $\sigma_{P'}$  (from each EOF reconstruction) (Sect. A6). All these quantities are obtained without gaps on the full spatial domain of the original model dataset  $\mathbf{X}$ .

Finally, due to the high temporal dynamics of our Baltic Sea environment, we use an expansion of the EOF reconstruction approach to reconstruct not only the data value, but both the data value and a (short-term) linear trend, in order to temporally collate (temporally extended) observations into a time-coherent, synoptic picture (Sect. A7; Elken et al., 2019).

### 2.2 Baltic Sea $p\text{CO}_2$ climatology

Given limitations of modelled data, we aimed to produce an observation-based  $p\text{CO}_2$  climatology. For this, we used



**Figure 1.** Approach to build a regional  $p\text{CO}_2$  climatology of the Baltic Sea. (1)  $p\text{CO}_2$  data from an ecosystem model are mathematically analysed for dominant patterns of variability in a truncated EOF analysis. (2) Patterns of variability and (3)  $p\text{CO}_2$  observations of a given month are used to reconstruct Baltic Sea  $p\text{CO}_2$  distributions  $\bar{\mathcal{X}}$ , which are combined through an ensemble approach into (4) a monthly climatology of surface  $p\text{CO}_2$ , including a  $p\text{CO}_2$  error estimate, and the (short-term)  $p\text{CO}_2$  trend, including the  $p\text{CO}_2$  trend error estimate. For illustration, time series are given for three sample locations for model  $p\text{CO}_2$  (green),  $p\text{CO}_2$  observations (black), and climatological  $p\text{CO}_2$  (red).

(a) the above extrapolation approach with spatial patterns based on ecosystem model data and (b) observations of surface  $p\text{CO}_2$  from the Surface Ocean  $\text{CO}_2$  Atlas (SOCAT) to produce a monthly climatology of  $p\text{CO}_2$  as well as of the linear, short-term  $p\text{CO}_2$  trend to temporally collate  $p\text{CO}_2$  observations (Sect. A7) within a month (see Fig. 1 for a visualization of the approach).

### 2.2.1 Spatial patterns of variability

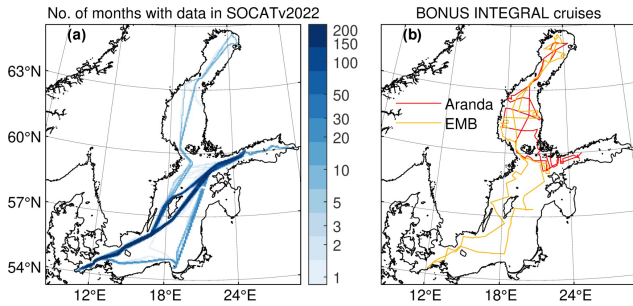
Spatial patterns of variability  $e_i$  for our extrapolation approach are based on model data from a Baltic Sea setup of the Ecological ReGional Ocean Model (ERGOM version 1.2). This version of ERGOM includes a simple carbon cycle as described in Kuznetsov and Neumann (2013) with amendments to allow for non-Redfield stoichiometry as outlined in Neumann et al. (2022). In principle, any carbon-containing biogeochemical ocean model of choice can be used as a basis for the patterns of variability  $e_i$ . Nonetheless, a better carbon

representation in the model likely provides better-suited  $e_i$  patterns of variability.

The model data variability characteristics are illustrated in the Appendix (Fig. A1). Less than 1% of the locations show a temporal decorrelation scale  $\leq 7$  d, so that we chose a weekly aggregation of the model data for variability pattern extraction. For the climatology itself, we chose a more typical monthly resolution.

ERGOM has been shown to adequately mirror observations of the large-scale nitrate, phosphate, oxygen, and carbon distribution (Neumann et al., 2022; Eilola et al., 2011; Neumann et al., 2015). However, shortcomings still exist in the exact magnitude and timing of phytoplankton carbon uptake and release throughout the seasons (Neumann et al., 2022; Fig. 4).

From an ERGOM version 1.2 model run from 1948 to 2020, we used the last 20 years of modelled surface  $p\text{CO}_2$  from 2001 to 2020 averaged into 1044 weekly means. The model run has a horizontal resolution of 3 nautical miles, which yields a dataset  $\mathbf{X}$  with 12 010 grid points  $m$  for the



**Figure 2.** (a) Map of the Baltic Sea with the number of individual months with  $p\text{CO}_2$  observations available in SOCATv2022 between June 2003 and December 2021 (maximum 223 months). (b) Cruise tracks of BONUS INTEGRAL cruises on RV *Aranda* in February and March 2019 and RV *Elisabeth Mann Borgese* in May and June 2019.

Baltic Sea area south of the Skaggeiak (i.e. HELCOM sub-basins 2 to 17, HELCOM Secretariat, 2017) and 1044 time steps  $n$ .

From these model data  $\mathbf{X}$ , patterns of variability  $e_i$  were extracted by EOF analysis (see Appendix A for details). The analysis of our model data retained 224 EOF modes with a minimal cross-validation error of  $22.6 \mu\text{atm}$  against the model data and a total explained variance of 98.6 %. The spatial patterns  $e_i$  have the same 3 nautical mile resolution as the model data.

### 2.2.2 Surface $p\text{CO}_2$ observations

Surface  $p\text{CO}_2$  measurements in the Baltic Sea were obtained from SOCAT version 2022 (Pfeil et al., 2013; Bakker et al., 2016; Bakker et al., 2022), which collects surface  $p\text{CO}_2$  data from underway observations of ships of opportunity (SOOPs) or research vessels. All observations are based on  $\text{CO}_2$  measurements in air equilibrated with sea surface waters (Körtzinger et al., 1996; Pierrot et al., 2009) with a typical accuracy of 2–5  $\mu\text{atm}$  (in some cases  $\leq 10 \mu\text{atm}$ ), which is indicated by the respective quality flags A–E (Pierrot et al., 2009; Bakker et al., 2016).

Here, Baltic Sea  $p\text{CO}_2$  data for the period June 2003 to December 2021 were used. They were from the ICOS-DE SOOP *Finnpartner/Finnmaid* line (Schneider et al., 2006; Güllow et al., 2011) between Lübeck-Travemünde and Helsinki, which covers the southern and central Baltic Sea, or the ICOS-SE SOOP *Tavastland* line between Lübeck-Travemünde and Oulu or Kemi, which additionally covers the northern basins starting from 2019 (Fig. 2a). Additional surface  $p\text{CO}_2$  data originated from RV *Aranda* cruise ARA04/2019 in February and March 2019 and RV *Elisabeth Mann Borgese* cruise EMB214 (Rehder et al., 2021) in May and June 2019 (Fig. 2b), which were carried out as part of the BONUS INTEGRAL project. Data processing and quality control followed the SOCAT guidelines. In total, data cover-

age in the southern and central Baltic Sea is high, with locally up to 189 out of 223 months covered (i.e. June 2003 to December 2021). Data in the northern Baltic Sea, however, are only available since February 2019, with observations during a total of 15 out of 35 months.

### 2.2.3 Monthly climatology construction

For every month  $t$  of the 189 months with observations, the observations were centred temporally on the 15th of each month as  $t_o$  (see Sect. A7), so that the extrapolation approach provides a field  $\bar{\mathcal{X}}$  with a spatial distribution of

- the reconstructed  $p\text{CO}_2$  ( $x_{t,\text{reconstr}}$ , Eq. A27) at  $t_o$ ,
- the  $p\text{CO}_2$  error estimate ( $\sigma_{t,\text{reconstr}}$ , Eq. A28) at  $t_o$ ,
- the (short-term)  $p\text{CO}_2$  trend at  $t_o$ ,
- an error estimate of the  $p\text{CO}_2$  trend at  $t_o$ , and
- the average number of patterns  $e_i$  used in the reconstruction ( $\bar{l}$ , Eq. A25)

for each month with observations (e.g. Fig. 3 for  $\bar{\mathcal{X}}$ s of May 2013 as an illustration).

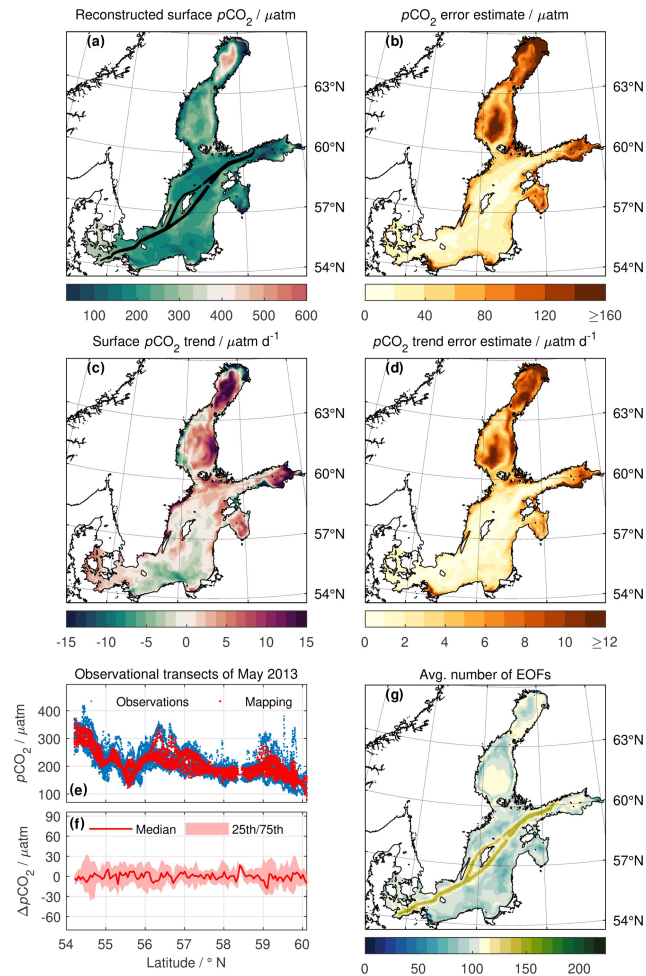
The construction of a monthly  $p\text{CO}_2$  climatology for the entire Baltic Sea has its bottleneck in the temporal coverage of the northern basins, i.e. the Bothnian Sea, the Quark, and the Bothnian Bay. While the central and southern Baltic Sea has approximately 85 % of all the months covered since June 2003, observations in the northern basins only start in February 2019, with 43 % monthly coverage until the end of 2021. While the extrapolation approach provides fully filled maps of  $p\text{CO}_2$ , including the northern basins, also for months pre-February 2019, they are associated with larger uncertainties, i.e. limited explanatory power, in the northern basins (e.g. Fig. 3b). To not create a wrong impression of seasonality or interannual variability, especially if the mapping uncertainty fields  $\sigma_{t,\text{reconstr}}$  are not respected by a casual user, we chose to provide a mean seasonal climatology  $\mathbf{Y}$ . Here, the weighting of the climatology (Eq. 1; see below) ensures the preference of reconstructions with proper observation constraints (i.e. February 2019 and onwards for the northern basins) over reconstructions without them (i.e. pre-February 2019).

For a given location  $m$ ,  $\chi_m$  denotes the (reconstructed) time series of  $\bar{\mathcal{X}}$  (with size  $1 \times 189$ ) and  $y_m$  the time series of the mean monthly climatology  $\mathbf{Y}$  (with size  $1 \times 12$ ), respectively. The monthly means  $y_m$  were then obtained by

$$\mathbf{w}_m \cdot \chi_m = \mathbf{w}_t \cdot \left( \mathbf{O}_m^T \mathbf{y}_m^T + \mathbf{g}_m \cdot (t - t_{\text{ref}}) \right) \quad \text{with} \quad (1)$$

$$\mathbf{w}_m = 1 / \sigma_{m,\text{reconstr}}^2, \quad (2)$$

where  $\mathbf{w}_m$  is the time series of inverse variance weights (Eq. 2) with size  $1 \times 189$  at the given location  $m$ ,  $\mathbf{O}_m$  a time



**Figure 3.** Monthly fields  $\bar{\mathcal{X}}$  of surface  $p\text{CO}_2$  reconstruction for May 2013. (a) Reconstructed  $p\text{CO}_2$   $x_{t,\text{reconstr}}$  with observation locations given as black dots. (b)  $p\text{CO}_2$  error estimate  $\sigma_{t,\text{reconstr}}$  with small values close to observations. (c)  $p\text{CO}_2$  trend and (d) its error estimate. (e) Observed (blue) and reconstructed  $p\text{CO}_2$  (red) against latitude as well as (f) their difference as the median (solid line) and interquartile range (pale shading). (g) Average number of patterns  $\bar{l}$  of the reconstruction: above-average  $\bar{l}$  indicates the use of high-order, smaller-scale patterns  $e_i$ , notably close to data constraints from observations, whereas below-average  $\bar{l}$  indicates areas where low-order, larger-scale patterns  $e_i$  dominate the reconstruction ensemble (see Eq. A26). Compare Fig. A2 for a reconstruction without a temporal trend (Sect. A7).

series operator or matrix of size  $12 \times 189$  that selects the monthly mean  $y_m$  corresponding to the month of  $\chi_m$ ,  $t - t_{\text{ref}}$  the time difference (in decimal years) between each of the 189 monthly  $\bar{\mathcal{X}}$  and a reference time,  $t_{\text{ref}}$ , and  $g_m$  an extra degree of freedom to allow for a linear long-term trend in  $\bar{\mathcal{X}}$  for each location  $m$  (e.g. surface  $p\text{CO}_2$ ). As a reference time, the middle of 2013 was chosen, which closely corresponds to the mean date of all the observations.

Equation (1) represents a system of 189 linear equations with 13 degrees of freedom to calculate 12 weighted mean values  $y_m$  and a long-term trend  $g_m$ . Through the inverse variance weights (Eq. 2), monthly maps  $\chi_m$  with better-constrained data, i.e. smaller variance of reconstruction  $\sigma_{m,\text{reconstr}}^2$  at the given location  $m$  (and for the given month), obtain preference in the weighted mean (Eq. 1).

When done for each point  $m$ , one obtains the mean seasonal climatology  $\mathbf{Y}$  (with size  $m \times 12$ ), centred on 2013, as well as a map of its mean linear (long-term) trend  $\mathbf{G}$  (with size  $m \times 1$ ) for each of the five fields  $\mathcal{X}$  discussed.

## 2.2.4 Dataset description

The dataset (Bittig et al., 2023) consists of two plain-text files. One contains the mean seasonal climatology  $\mathbf{Y}$  and the other the mean linear (long-term) trend  $\mathbf{G}$ . Based on the 3 nautical mile resolution of the original model data, both feature 12 010 grid points  $m$  for HELCOM sub-basins 2 to 17. For the months January through to December, the climatology file gives  $p\text{CO}_2$  ( $\mu\text{atm}$ ), a  $p\text{CO}_2$  error estimate ( $1\sigma$ ) ( $\mu\text{atm}$ ), a (short-term)  $p\text{CO}_2$  trend for the 15th of each month ( $\mu\text{atm d}^{-1}$ ), a (short-term)  $p\text{CO}_2$  trend error estimate ( $1\sigma$ ) ( $\mu\text{atm d}^{-1}$ ), and the average number of patterns ( $\bar{l}$ ) (see Figs. A3 and A4). The long-term trend file gives both the long-term linear trend  $\mathbf{G}$  ( $\mu\text{atm yr}^{-1}$ ) (Fig. 7) and its error estimate ( $1\sigma$ ) ( $\mu\text{atm yr}^{-1}$ ).

## 3 Results

### 3.1 Extrapolation approach

The mapping approach gives fully filled fields on the entire spatial domain from scattered observational data. The mapped  $p\text{CO}_2$  is in a similar range and frequency distribution to observations with improvements in terms of distribution and correlation over the original model  $p\text{CO}_2$  (Fig. 4a). The mapped fields are smooth and without spatial gaps or discontinuities. The  $p\text{CO}_2$  error estimate is reduced in the spatial vicinity of observations. In contrast,  $p\text{CO}_2$  error is markedly increased both near shore as well as in sub-basins not covered by observations, e.g. the northern basins or the Gulf of Riga for a SOOP *Finnmaid*-based reconstruction (e.g. Fig. 3b). Still,  $p\text{CO}_2$  values remain within reasonable margins even in those areas (e.g. Fig. 3a). We cannot observe a tendency of the mapping approach to give extreme values or outliers in the absence of observations (compare Fig. 4a). Rather, with few or no data constraints, variability pattern  $e_i$  amplitudes in the ensemble tend towards zero, i.e. towards zero deviation, which is equivalent to the spatially resolved temporal mean  $p\text{CO}_2$  of the model (Appendix A).

Addition of a linear temporal trend to collate observations to a common time  $t_0$  helps to reduce the mismatch between temporally spread observations and a full-domain mapping

of a dynamic coastal system (compare e.g. Figs. 3e, f and A2c, d, respectively).

However, for about half the grid points of the 189 monthly mappings, the magnitude of the short-term temporal  $p\text{CO}_2$  trend (e.g. Fig. 3c) at a given location  $m$  is insignificant, i.e. within its trend error estimate (e.g. Fig. 3d). At the same time, a considerable portion of the significant  $p\text{CO}_2$  trends is found at grid points  $m$  in the direct footprint of observations where the trend error estimates are small in the first place. This demonstrates that a monthly temporal resolution is sufficient for construction of a Baltic Sea  $p\text{CO}_2$  climatology with our approach.

### 3.2 Quality of interpolation and extrapolation

To assess the quality of the obtained fields, we consider (a) the mapped result against the original observations and (b) a comparison of mappings from concurrent subsets of observations.

For the first aspect, we consider the residual  $p\text{CO}_2$  against SOCAT observations for all 189 monthly mapped fields from the climatology construction (e.g. Fig. 3e, f). We use typical metrics for interpolation methods (correlation, bias, standard deviation). As the monthly fields  $\mathcal{X}$  were constructed from those observations, this assesses primarily to which degree the data constraints were taken into account for the mapping, e.g. the balance between point-to-point reproduction and smoothing with the given patterns of variability  $e_i$ .

The comparison shows highly correlated  $p\text{CO}_2$  data ( $R^2 = 0.95$ ), a similar  $p\text{CO}_2$  distribution (Fig. 4a), and an unbiased mapping with a standard deviation of  $26 \mu\text{atm}$  (Fig. 4b). This includes both misfits of the mapped fields as well as variations within observations of a given grid point location (compare the spread between observations and mapping in Fig. 3e). The error estimate is of a similar magnitude (Fig. 4c) and is well collocated. That is, residuals are within  $1 \times$  the error estimate in 68 % of the cases (Fig. 4d).

The second aspect, how data constraints in one area transfer to accurate predictions in another part of the domain, is more difficult to assess objectively. It requires the observational data to be split into a subset that is used for mapping and a subset that is deemed independent and used for evaluation. How this subsetting is done, however, directly influences the outcome of the evaluation. That is, while such evaluations may seem instructive and generalizable, they are highly dependent on design and thus carry some degree of arbitrariness.

To illustrate this, we consider a series of monthly mappings for May 2019, the month with the highest density of observations in our 189 monthly mappings. We use chessboard-like grids with alternating white and black boxes for subsetting, where observations in the white boxes are used for mapping and observations in the black boxes are used for evaluation. Grid box sizes vary from  $1^\circ \times 1^\circ$ ,  $2^\circ \times 2^\circ$ , to  $3^\circ \times 3^\circ$  subsetting grids (see Fig. 5a, e, and i

for example grids). To check for dependence on a specific data distribution, the  $1^\circ \times 1^\circ$ ,  $2^\circ \times 2^\circ$ , and  $3^\circ \times 3^\circ$  subsetting grids were shifted spatially in latitude and longitude so that every grid size had eight different realizations, each with different sets of data in the mapping and evaluation subsets, respectively (Fig. 5).

For  $1^\circ \times 1^\circ$  subsetting, the eight realizations show a relatively coherent picture: differences  $\Delta p\text{CO}_2$  between mapped and observed  $p\text{CO}_2$  (Fig. 5b) are slightly larger (by  $5\text{--}12 \mu\text{atm}$ ) than for interpolated data (Fig. 4b) and are mostly unbiased.  $p\text{CO}_2$  error estimates are elevated (Fig. 5c vs. Fig. 4c), while the ratio of differences and error estimates, i.e. the collocation of bias and error estimate, shows a similar distribution for both extrapolation (Fig. 5d) and interpolation (Fig. 4d). With increasing spatial scales of the  $2^\circ \times 2^\circ$  and  $3^\circ \times 3^\circ$  subsetting,  $p\text{CO}_2$  differences and error estimates become larger, as would be expected (Fig. 5f, j and g, k). Encouragingly, however, the ratio of differences and error estimates preserves a similar shape (Fig. 5h, l). That is, the mapping approach seems to provide suitable mapping error estimates.

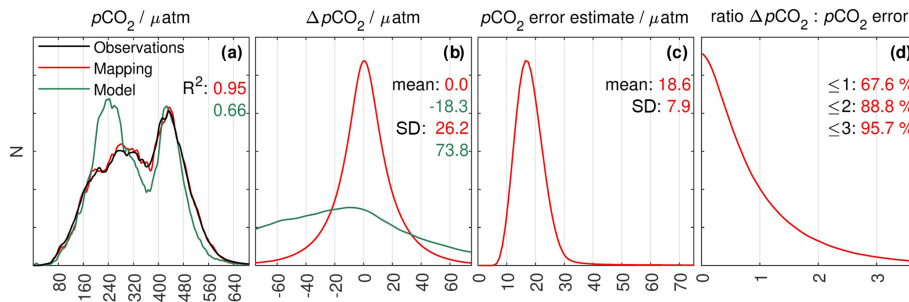
More concerning, however, are increasing differences between the eight realizations with increasing spatial scales. Especially for the  $3^\circ \times 3^\circ$  subsetting, some realizations show a  $p\text{CO}_2$  bias, which additionally can be different between realizations (blue vs. orange realizations, Fig. 5j), as well as noticeable differences in  $p\text{CO}_2$  error estimates and thus mapped  $p\text{CO}_2$  (e.g. green realization, Fig. 5k). This illustrates a strong dependence on the specific choice of subsetting, i.e. which data are included in mapping and which data are excluded for validation.

We conclude that, with a given choice of subsetting, one implicitly chooses which statistical values one wants to get out. While our evaluation could be described as being well designed, it turns out to be starting-point-dependent. If a month with less data coverage were chosen, differences between realizations (i.e. which data end up in the training or validation subset) would be even larger. Therefore, based on the above illustration, any subsetting-based evaluation needs to be taken with a sufficient amount of caution, as the subsetting design will imprint itself on the outcome.

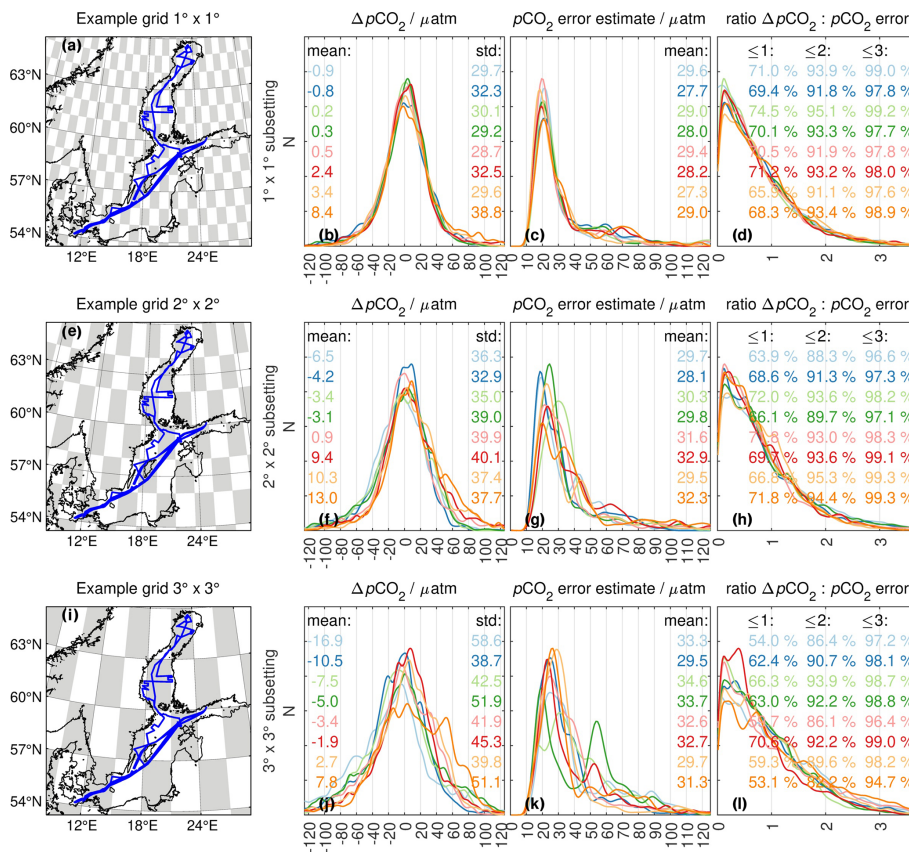
## 3.3 Baltic Sea $p\text{CO}_2$ climatology

### 3.3.1 $p\text{CO}_2$ distribution

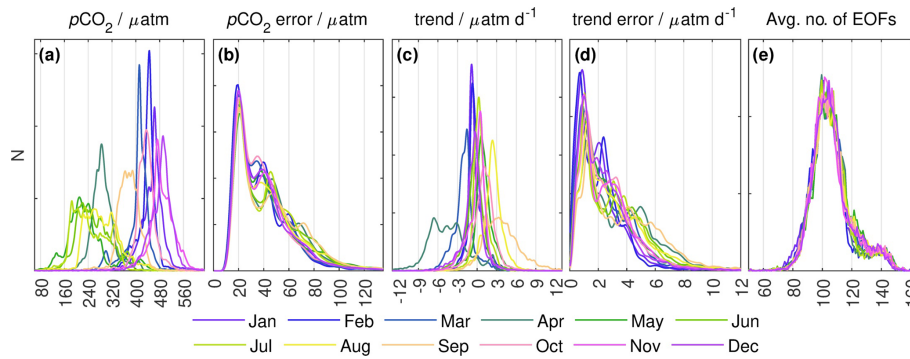
The mean seasonal  $p\text{CO}_2$  distribution  $\mathbf{x}_{\text{reconst}}$  matches previously published seasonal cycles of surface  $p\text{CO}_2$  in the western and central Baltic Sea (Schneider and Müller, 2018): a strong spring bloom  $p\text{CO}_2$  drawdown occurs between March and May to a low around  $200 \mu\text{atm}$  in May, some slight increase in June, and a second summer bloom low in July, followed by relaxation towards atmospheric levels in September and October and supersaturated levels during late autumn



**Figure 4.** Histograms for (a)  $p\text{CO}_2$  observations (black) as described in Sect. 2.2.2,  $p\text{CO}_2$  reconstructed at the observation times and locations from the 189 monthly mappings (red), and model  $p\text{CO}_2$  extracted at the observation times and locations (green). Reconstructed and observed  $p\text{CO}_2$  data are highly correlated ( $R^2$ ), and they show a similar distribution. Correlation between modelled and observed  $p\text{CO}_2$  is weaker, and their distributions show noticeable differences. (b) The  $p\text{CO}_2$  difference between reconstructed (red) or modelled (green)  $p\text{CO}_2$  and observed  $p\text{CO}_2$ . There is no bias in the reconstruction, while differences between model and observations show a much wider range than the reconstructed  $p\text{CO}_2$ . (c) The  $p\text{CO}_2$  error estimates  $\sigma_{\text{reconstr}}$  of the reconstruction at the observation times and locations. (d) The ratio between the  $p\text{CO}_2$  difference and  $p\text{CO}_2$  error estimate, where a ratio  $\leq 1$  means that the observed  $p\text{CO}_2$  is within  $1 \times \sigma_{\text{reconstr}}$  of the mapped  $p\text{CO}_2$ .



**Figure 5.** Subsetting evaluations for May 2019 observations using  $1^\circ \times 1^\circ$  subsets (top row; a–d),  $2^\circ \times 2^\circ$  subsets (middle row; e–h), and  $3^\circ \times 3^\circ$  subsets (bottom row; i–l), respectively, with maps (a, e, i) illustrating one example subsetting grid where observations (blue) in white boxes are used for  $p\text{CO}_2$  reconstruction and observations in grey boxes for evaluation. Histograms are for (b, f, j) the  $p\text{CO}_2$  difference between reconstructed and observed  $p\text{CO}_2$  within the evaluation boxes. (c, g, k) The  $p\text{CO}_2$  error estimate  $\sigma_{\text{reconstr}}$  of the reconstruction at the observation times and locations within the evaluation boxes. (d, h, l) The ratio between the  $p\text{CO}_2$  difference and  $p\text{CO}_2$  error estimate within the evaluation boxes, where a ratio  $\leq 1$  means that the observed  $p\text{CO}_2$  is within  $1 \times \sigma_{\text{reconstr}}$  of the mapped  $p\text{CO}_2$ . The colours represent eight different  $1^\circ \times 1^\circ$ ,  $2^\circ \times 2^\circ$ , and  $3^\circ \times 3^\circ$  subsetting grids only offset spatially in latitude and longitude.



**Figure 6.** Smoothed histograms for the five output fields of the climatology  $\mathbf{Y}$ : (a) reconstructed  $p\text{CO}_2$   $x_{t,\text{reconstr}}$ , (b)  $p\text{CO}_2$  error estimate  $\sigma_{t,\text{reconstr}}$ , (c)  $p\text{CO}_2$  short-term trend, (d) error estimate in the  $p\text{CO}_2$  short-term trend, and (e) the average number of patterns  $\bar{l}$  used in the reconstruction. The colour code indicates the respective month.  $p\text{CO}_2$  follows a typical seasonal cycle, while the  $p\text{CO}_2$  error estimate and the number of patterns show no apparent seasonal variations.  $p\text{CO}_2$  trends show higher magnitudes during spring (March and April) and late summer (August and September), which is mirrored by enhanced  $p\text{CO}_2$  trend error estimates in these months.

and winter, peaking around  $500 \mu\text{atm}$  in December (Fig. 6a; Figs. A3 and A4, first column).

Regionally, the western basins lead the mean seasonal cycle, while the central Gotland Basin and the Gulf of Finland trail behind. The productive season is even shorter in the northern basins, with the major  $p\text{CO}_2$  drawdown in the Bothnian Sea occurring in June (Figs. A3 and A4, first column). Similarly, the seasonal amplitude is less pronounced in the western or northern basins than in the Gotland Basin and is most intense in the Gulf of Finland. For the northern basins, this is the first fully seasonal climatology from  $p\text{CO}_2$  observations.

The  $p\text{CO}_2$  error estimate  $\sigma_{\text{reconstr}}$  shows a minimum of  $12 \mu\text{atm}$ . Close to observations, the  $p\text{CO}_2$  error estimate of the mean seasonal climatology is of the order of  $20 \mu\text{atm}$  and as such close to the value of the individual mappings (compare Fig. 6b and Fig. 4c). Further away,  $\sigma_{\text{reconstr}}$  increases up to a level around  $90 \mu\text{atm}$  (95th percentile), with a spatial preference for high  $\sigma_{\text{reconstr}}$  in the northern basins as well as in the Gulf of Riga, the Oder Bight, and other sheltered, near-shore areas (Figs. A3 and A4, second column).

### 3.3.2 $p\text{CO}_2$ short-term trends

In general, the short-term  $p\text{CO}_2$  trends follow the mean seasonal  $p\text{CO}_2$  cycle with on average negative trends, i.e. decreasing  $p\text{CO}_2$ , during winter and spring. Positive trends prevail during late summer and autumn (Fig. 6c), with the strongest trend magnitudes occurring in the northern Baltic proper and the Gulf of Finland (Figs. A3 and A4, third column).

However, in 8 out of the 12 months, the majority of estimated trends are insignificant. That is, more than 70% are smaller than their error estimate, indicating that inclusion of a short-term trend in the mapping may not be required for these months. Only in March and April as well as in Au-

gust and September are the majority of  $p\text{CO}_2$  trends significant (despite elevated  $p\text{CO}_2$  trend error estimates occurring at the same time in some areas; see Fig. 6d). That is, significant short-term  $p\text{CO}_2$  trend magnitudes fall together with the strong springtime  $p\text{CO}_2$  drawdown as well as with autumnal mixed-layer deepening and entrainment of high- $p\text{CO}_2$  waters (Fig. 6; more in Jacobs et al., 2021).

Like for the  $p\text{CO}_2$  distribution, trend error estimates are higher in the northern basins as well as the Gulf of Riga, Oder Bight, and other sheltered, near-shore areas (Figs. A3 and A4, fourth column).

### 3.3.3 Number of patterns or mapping scales

The average number of patterns  $\bar{l}$  per month can be seen as a qualitative indicator to show preference for reconstructions with larger or smaller spatial scales, respectively, in the weighted mean (Eq. A25) of the reconstruction.

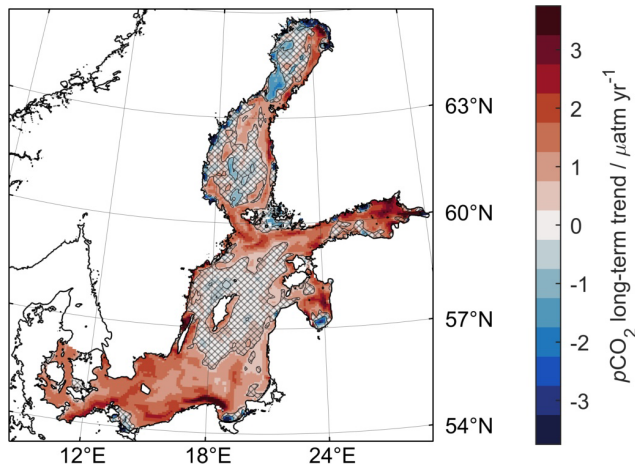
For the climatology  $\mathbf{Y}$ ,  $\bar{l}$  shows a maximum around 105 patterns (Fig. 6e). This level is at about 45% of the maximum number of patterns, i.e. indicating only a slight preference for reconstructions with larger-than-average scales in the weighted mean. However, a notable fraction with a higher number of patterns  $\bar{l}$  and a peak around 140 patterns exists (Fig. 6e). These are located in the vicinity of the observations, where the preference in the weighted mean is for reconstructions with stronger small-scale features (see e.g. Fig. 3g).

There is no seasonal imprint on the average number of patterns  $\bar{l}$ , which is in contrast to the  $p\text{CO}_2$  and  $p\text{CO}_2$  short-term trend distributions (Fig. 6).

### 3.4 Long-term $p\text{CO}_2$ trend

The mean climatology  $\mathbf{Y}$  is constructed from the 189 monthly mappings  $\mathbf{X}$  and a linear long-term trend  $\mathbf{G}$  (see Eq. 1 and Sect. 2.2.3). For  $p\text{CO}_2$ , the analysis shows a signif-





**Figure 7.** Long-term trend  $G$  of surface  $p\text{CO}_2$  from the monthly mappings and the climatology construction (Eq. 1). Areas where the long-term trend is not significant are cross-hatched. In the southern Baltic Sea, a significant increase of  $+1.4 \mu\text{atm yr}^{-1}$  is obtained from the observations. Selected parts of the northern Baltic proper, the Åland Sea, and the Gulf of Finland give a similar increase of  $+1.5 \mu\text{atm yr}^{-1}$ , while the western and eastern Gotland Basin show no significant long-term trend.

icant long-term  $p\text{CO}_2$  increase of  $+1.4 \pm 0.5 \mu\text{atm yr}^{-1}$  ( $1\sigma$ ) in the southern Baltic Sea from the Great Belt to the Bornholm Basin and southern parts of the Gotland Basin (Fig. 7). A similar increase ( $+1.5 \pm 0.7 \mu\text{atm yr}^{-1}$ ) is obtained in selected regions of the northern Baltic proper, the Åland Sea, and the Gulf of Finland. A positive trend in  $p\text{CO}_2$  is also visible in coastal areas along the Finnish coast in the northern basins as well as in the Gulf of Riga. However, multi-year observations in those areas are unavailable to support a long-term  $p\text{CO}_2$  trend (see Sect. 2.2.2; Fig. 2). In all the other areas, the data available do not indicate a significant long-term trend, i.e. one that is outside the 68.3% ( $1\sigma$ ) confidence bound. This includes in particular the western and eastern Gotland Basin but also the northern basins (Fig. 7), where the observational time series is likely too short to allow for the detection of a significant long-term trend (Fig. 2).

## 4 Discussion

### 4.1 Extrapolation approach

The extrapolation approach uses an EOF analysis of the data covariance  $\mathbf{Q}$  as a foundation. When working with an EOF reconstruction, one needs to bear in mind the purely mathematical nature of the EOF modes'  $e_i$  construction. They are built by (1) maximizing the (partial) variance they explain and by (2) being orthogonal to all previous EOF modes (Preisendorfer, 1988; Monahan et al., 2009; Dommenges, 2015). They may therefore indicate some causal connection within the original data  $\mathbf{X}$ , which requires confirmation by

other means. They may, however, just as well group some portion of (partial) variance in one part of the domain with some portion of (partial) variance in another part of the domain just to mathematically maximize the amount of (partial) variance of the given EOF mode  $e_i$ . This can give rise to apparent EOF-mode “teleconnections”, where one part of the data or spatial domain is seemingly tightly coupled to another part of the data, e.g. for the second EOF mode  $e_{i=2}$ , which always has a dipole pattern. Often, there is a temptation to attribute such couplings or “teleconnections” to a driving mechanism. However, given its origin, patterns of EOF modes should be interpreted with great caution (and only with ancillary, supporting data) and are best seen to be of a mathematical rather than mechanistic nature (Monahan et al., 2009; Dommenges, 2015). We therefore did not try to assign physical mechanisms or drivers to specific EOF patterns  $e_i$  in our work.

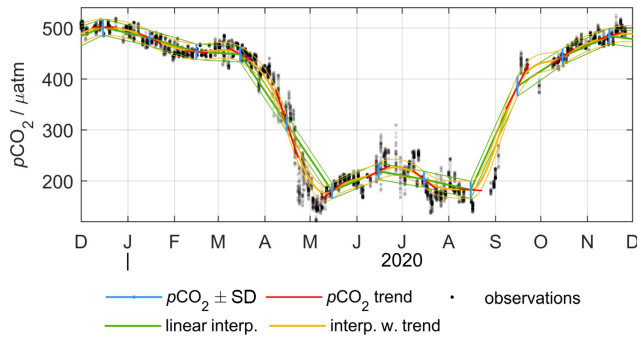
The use of EOF modes as a basis for extrapolation ensures that the extrapolated map covers the full spatial domain, that it is gap-free, and that it is discontinuity-free.

A key aspect of truncated EOF reconstructions is the number  $l$  of significant modes included. Our cross-validation approach provided a relatively large number of  $l_{\text{max}} = 224$  modes. The relatively large number of modes is a prerequisite for a small representational error  $\mathbf{P}'$ , i.e. for the covariance not resolved by the collection of truncated EOF modes, which gives a lower bound on the uncertainty of the mapping  $\sigma_{\text{reconstr}}$  (Eq. A28). In our case, the average  $\sigma_{P'}$  when using all 224 patterns  $e_i$  amounts to  $14.7 \mu\text{atm}$ . If only the first 50 patterns  $e_i$  were included, the average representational error would be as high as  $29.1 \mu\text{atm}$ .

A reconstruction with a small number of modes provides for a more uniform, large-scale homogeneous field, where gaps in the data are filled by the large-scale picture. However, such reconstructions may lack the flexibility to reproduce real features of the observations, e.g. through too strong smoothing. Conversely, a large number of modes provides for a fine, small-scale field with high flexibility. However, features in some areas without nearby observations may be badly constrained with the risk of “ghost” signals.

By using an ensemble of reconstructions that cover the entire range of  $1 \dots l_{\text{max}}$  modes and by using the mapping variance  $\sigma_P^2$  as weights (Eqs. A25 and A26), we find an optimal trade-off in our mapping approach, where both aspects are balanced. In this way, the scales used in the mapping are adjusted spatially according to the distribution of observations as constraints.

The cost function of the mapping approach minimizes the residual between observations and mapped data (Eq. A19). In this way, a synoptic  $p\text{CO}_2$  mapping is obtained. If observations, however, are temporally extended compared to the system's temporal dynamics, the  $p\text{CO}_2$  mapping can be distorted and cannot represent a synoptic picture, depending on the spatial and temporal pattern by which observations are obtained (Elken et al., 2019).



**Figure 8.** Difference between monthly point-by-point interpolation (green) or interpolation with a trend (orange) in a dynamic coastal system. Example data are from the northern Baltic proper (approximately  $58^{\circ}45' \text{ N}$ ,  $021^{\circ}00' \text{ E}$ ), with  $p\text{CO}_2$  observations  $\mathbf{X}$  shown in black. They are the basis for monthly mapped fields  $\bar{\mathbf{X}}$  of surface  $p\text{CO}_2$  (blue) as well as for a linear (short-term)  $p\text{CO}_2$  trend (red). Compared to a linear interpolation between monthly  $p\text{CO}_2$  values (green), interpolation including the trend information (orange) gives a better reproduction of the observed dynamics, particularly in spring and autumn. A cubic Hermite spline was used for the calculation (Appendix B).

The introduction of a temporal trend at each location (Sect. A7) as an additional degree of freedom reduces such distortion caused by sampling. It both improves the fit to the data (compare Fig. A2c, d vs. Fig. 3e, f) and provides extra information.

- For a given mapping, a strong or weak trend indicates where temporal dynamics are high or low and informs on where the frequency of observations should be enhanced or not, respectively.
- For a time series of mappings, information on both the value and trend allows for a more accurate interpolation by a cubic Hermite spline (Appendix B) compared to a standard, linear interpolation of the values alone (e.g. Fig. 8).

For our mapping of surface  $p\text{CO}_2$  in monthly intervals, we conclude that the inclusion of a temporal trend in the mapping is required, particularly for spring and autumn. For months without strong trends, we do not see adverse effects in the mapped  $p\text{CO}_2$  compared to an extrapolation without inclusion of a temporal trend.

The dynamics of the  $p\text{CO}_2$  field are prescribed by the model data covariance matrix  $\mathbf{Q}$  (Fig. A1a). The closer the model and real-world variability match, the easier it is for the mapping to reflect reality and to provide a most realistic picture. Nonetheless, also for a model dataset  $\mathbf{X}$  that may lack some real-world features or their magnitude, the observational constraints provide an improvement in the mapped field compared to the original model (Fig. 4).

The mapping error estimates are elevated where observations are scarce or dynamics are high to start with. For our

$p\text{CO}_2$  mapping, the model covariance  $\mathbf{Q}$  is the origin of a pronounced coast–basin difference in the  $p\text{CO}_2$  variance caused by both physical and biological drivers in the model. This prescribed data covariance distribution is retained in the mapping error estimates  $\sigma_{\text{reconstr}}$ . For example, they are enlarged in dynamic river plumes, in the sheltered, shallow Gulf of Riga, or in other near-shore areas, unless there are observations to constrain dynamics and thus reduce  $\sigma_{\text{reconstr}}$ . In regions with few observations,  $\sigma_{\text{reconstr}}$  approaches the prescribed data variance from  $\mathbf{Q}$  (Fig. A1a).

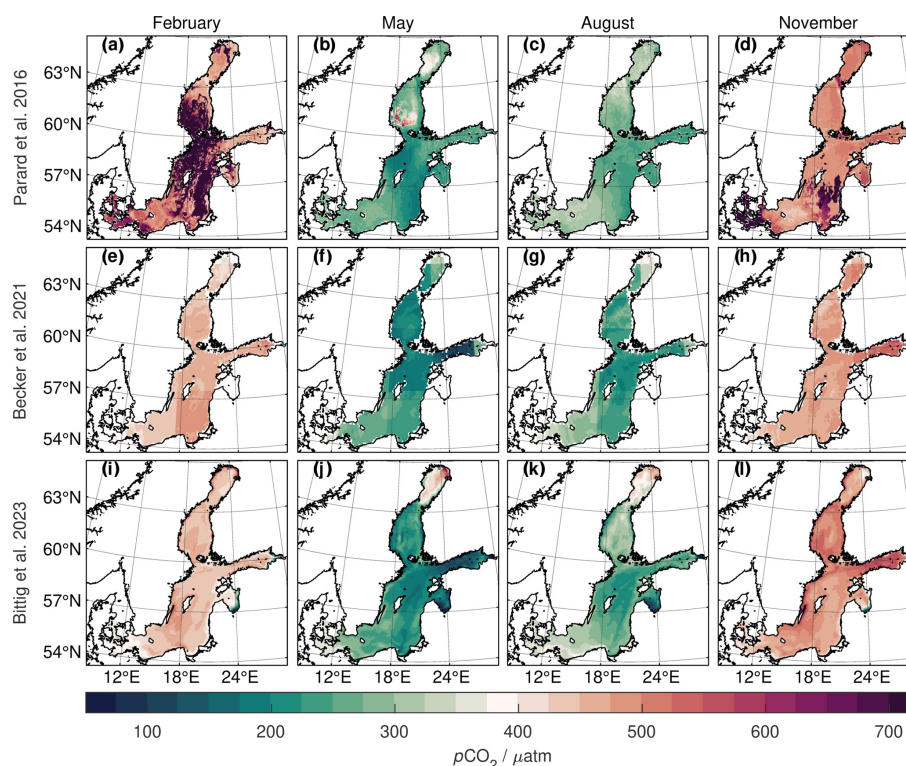
We do not observe a seasonality in the number of patterns  $\bar{l}$  of the mapping (Fig. 6e). At the same time, the  $p\text{CO}_2$  value varies strongly between seasons. In addition, there is a slight preference for observations to occur in spring, summer, and autumn over winter months. However, there is no discernible difference in the spatial coverage of  $p\text{CO}_2$  observations between seasons (data not shown). The aseasonality of  $\bar{l}$  can be explained by the “large-scale” mapping error  $\sigma_P$ , used for the ensemble weights (Eq. A26) to only depend on the distribution of samples  $\mathbf{H}_t$  and their respective observational error estimates  $\mathbf{S}$ . Therefore, the extrapolation approach does not discriminate or distinguish by ( $p\text{CO}_2$ ) value, and there is no seasonal imprint in the number of patterns.

#### 4.2 Comparison with other $p\text{CO}_2$ mapping approaches

SOCAT observations have been used previously to build surface  $p\text{CO}_2$  climatologies of the Baltic Sea (see Fig. 9 for examples). They rely on observation-driven extrapolation approaches, which can be grouped into different categories: (a) one where sparse data points are extrapolated based on some statistic metric of influence, e.g. a spatial decorrelation length or timescale, and (b) one where parametrizations with proxy data in combination with filled fields of those proxies are used for extrapolation (Rödenbeck et al., 2015).

Our mapping approach belongs to the first category. For a given month or time window, we use the available  $p\text{CO}_2$  data directly and one to one without an added transformation to inform the extrapolated  $p\text{CO}_2$  map. As a (spatial) metric of influence, the variability patterns are used. This gives a direct link between observations of a given month or time window and the mapped field. Conversely, if there are no  $p\text{CO}_2$  observations in a given time window, this directly translates to a (time) gap in the mapped fields for any approach of the first category.

Most previous  $p\text{CO}_2$  mappings in the Baltic Sea belong to the second category (e.g. Parard et al., 2016; Becker et al., 2021). In a first step, they use all available  $p\text{CO}_2$  data together with collocated proxies (e.g. location, sea surface temperature, chlorophyll-*a* concentration, water depth, distance to the coast, season) to establish a relationship or parametrization between proxies and  $p\text{CO}_2$ . In a second step, distributions of those proxy parameters are used to establish a  $p\text{CO}_2$  distribution in time and space. Here,  $p\text{CO}_2$  data of a given month do *not* directly inform the extrapolated  $p\text{CO}_2$



**Figure 9.** Mean monthly  $p\text{CO}_2$  distribution for 1 month per season (left to right) for different climatologies: Parard et al. (2016) (a–d), Becker et al. (2021) (e–h), and this work (i–l). The first two climatologies are proxy-based, where artefacts of the proxy directly translate to  $p\text{CO}_2$ , e.g. for wintertime satellite retrievals (a, d) or the subtle  $5^\circ \times 4^\circ$  gridding in Becker et al. (2021) (e–h). Our approach uses observed  $p\text{CO}_2$  data directly for mapping, which gives fewer artefacts but is only applicable when  $p\text{CO}_2$  observations are available.

map but only indirectly through the established relationship using all the data. Conversely, if there are no  $p\text{CO}_2$  observations in a given time window but proxy data are available, they can be used to estimate a  $p\text{CO}_2$  map nonetheless. That is, (temporal) gap filling is possible with the two-step procedure of approaches of the second category.

As a drawback, mapped  $p\text{CO}_2$  fields of the second category rely on the validity of the proxy relationship and, equally important, on good proxy input data. In some cases, e.g. satellite-based sea surface temperature or chlorophyll- $a$  concentration, proxies may be (seasonally) prone to artefacts, e.g. due to stronger cloud cover in winter months. Such artefacts then directly translate into mapped  $p\text{CO}_2$  fields and may create unrealistic patterns or out-of-range values (e.g. Fig. 9a, d; Parard et al., 2016). A mapping approach of the first category such as the one presented here, which uses the observations directly, is not affected by external-source data quality (Fig. 9i–l).

### 4.3 Long-term $p\text{CO}_2$ trend

The long-term trend of surface  $p\text{CO}_2$  in Baltic Sea surface waters has been studied before (Table 1): Wesslander et al. (2010) describe an increase in monthly  $p\text{CO}_2$  (approximately  $+20 \mu\text{atm yr}^{-1}$ ) for the 1993–1998 period in the eastern Got-

land Basin, which seems mostly driven by less intense summer  $p\text{CO}_2$  minima (from approximately  $50 \mu\text{atm}$  in 1993 to approximately  $250 \mu\text{atm}$  in 1998). For the longer period 1993–2009, they do not detect a significant trend. While Wesslander et al. (2010) use  $p\text{CO}_2$  calculated from pH and alkalinity, all further studies are based on the same surface  $p\text{CO}_2$  observations archived in SOCAT (Bakker et al., 2016), albeit using different methods and time extents.

Laruelle et al. (2018) found an increase of  $+2.9 \pm 2.4 \mu\text{atm yr}^{-1}$  ( $1\sigma$ ) for wintertime  $p\text{CO}_2$  from 1995 to 2011 in the southern and central Baltic Sea. At the same time, they see a similar spatial gradient to our work (Fig. 7) with a stronger increase in the south-west and a zero or even negative trend in the northern Gotland Basin.

Schneider and Müller (2018) describe an increase of  $+4.6$  to  $+6.1 \mu\text{atm yr}^{-1}$  ( $1\sigma$  between  $0.6$  and  $1.5 \mu\text{atm yr}^{-1}$ ) for different areas for the 2008–2015 period. However, their analysis did not deseasonalize the observations prior to linear trend analysis. During their analysed period, there is an increase in observations during autumn and winter (with typically high  $p\text{CO}_2$ ) and a decrease in early spring observations (with typically low  $p\text{CO}_2$ ), while late-spring and summer observations have a similar data coverage over the years. Their relatively high  $p\text{CO}_2$  trend estimate therefore seems to be

**Table 1.** Summary of long-term  $p\text{CO}_2$  trends for different areas and time periods in the Baltic Sea as obtained from different studies.

Study	Time period	Trend $\pm 1\sigma$ ( $\mu\text{atm yr}^{-1}$ )	Area	Note
Wesslander et al. (2010)	1993–1998	+20	Eastern Gotland Basin	$p\text{CO}_2$ calculated from pH and alkalinity
Wesslander et al. (2010)	1993–2009	Not significant	Eastern Gotland Basin	$p\text{CO}_2$ calculated from pH and alkalinity
Laruelle et al. (2018)	1995–2011	$+2.9 \pm 2.4$	Southern and central Baltic Sea	On wintertime $p\text{CO}_2$
Schneider and Müller (2018)	2008–2015	$+4.6$ to $+6.1$ ( $1\sigma$ 0.6–1.5)	Southern and central Baltic Sea	Shift in seasonal data coverage not accounted for in trend analysis
Becker et al. (2021)	1995–2016	$+2.05 \pm 0.12$	Southern Baltic Sea (south of $56^\circ\text{N}$ )	
Becker et al. (2021)	1995–2016	$+1.84 \pm 0.21$	Central Baltic Sea ( $56$ – $61^\circ\text{N}$ )	
This work (Fig. 7)	2003–2021	$+1.4 \pm 0.5$	Southern Baltic Sea (south of $56.5^\circ\text{N}$ )	
This work (Fig. 7)	2003–2021	Not significant	Central Baltic Sea; northern basins	
This work (Fig. 7)	2003–2021	$+1.5 \pm 0.7$	Åland Sea, northern Baltic proper, Gulf of Finland ( $59$ – $61^\circ\text{N}$ )	
This work (Fig. 7)	2003–2021	$+1.2 \pm 0.6$	Coastal waters in the northern basins and Gulf of Riga	Without direct observational support

partially caused by a seasonal shift in data availability that was left unaccounted for in the analysis.

Becker et al. (2021) derived trends of  $+2.05 \pm 0.12 \mu\text{atm yr}^{-1}$  for the southern Baltic Sea (south of  $56^\circ\text{N}$ ) and  $+1.84 \pm 0.21 \mu\text{atm yr}^{-1}$  for the central Baltic Sea (between  $56$  and  $61^\circ\text{N}$ ) for the period 1995–2016. However, they observe a stronger trend in earlier years (starting in the 1990s) than for later periods (starting in the mid-2000s).

Our analysis covers the more recent period 2003–2021 and gives significant trends of  $+1.4 \pm 0.5 \mu\text{atm yr}^{-1}$  ( $1\sigma$ ) in the southern Baltic Sea (south of  $56.5^\circ\text{N}$ ) and  $+1.5 \pm 0.7 \mu\text{atm yr}^{-1}$  in the Åland Sea, parts of the northern Baltic proper, as well as the Gulf of Finland (between  $59$  and  $61^\circ\text{N}$ ). The central Baltic Sea and the northern basins themselves show no significant trend. The positive trend ( $+1.2 \pm 0.6 \mu\text{atm yr}^{-1}$ ) in coastal waters of the northern basins along the Finnish coast as well as in the Gulf of Riga (Fig. 7) currently has no support from direct observational constraints in those areas (Fig. 2). In contrast, it likely originates from high similarity to other coastal waters (e.g. along the southern Baltic shore) in the model data variability analysis and thus in the variability patterns  $e_i$  used for mapping. As such, the  $p\text{CO}_2$  increase may be realistic but lacks field data support.

Together with the literature, our results seem to indicate a reduction in overall surface  $p\text{CO}_2$  increase in the Baltic Sea during the past decades or even its complete absence like in the central Baltic Sea (Fig. 7) in recent periods. While atmospheric  $p\text{CO}_2$  levels are still rising (approximately  $+2.4 \mu\text{atm yr}^{-1}$  in the past decade; Lan et al., 2022), this has implications for the  $\text{CO}_2$  balance between the atmosphere and the Baltic Sea. This supports previous claims of coastal seas becoming a less intense source of  $\text{CO}_2$  or eventually turning into (an increasing)  $\text{CO}_2$  sink (Laruelle et al., 2018).

## 5 Data availability

Surface  $p\text{CO}_2$  data are available from SOCAT (<https://doi.org/10.25921/1h9f-nb73>, Bakker et al., 2022; Bakker et al., 2016) and Rehder et al. (2021). ERGOM model output data are available at [https://thredds-iow.io-warnemuende.de/thredds/catalogs/projects/integral/catalog\\_pocNP\\_V04R25\\_3nm\\_agg\\_time.html?dataset=IOW-THREDDS-Baltic\\_pocNP\\_V04R25\\_3nm\\_agg\\_time\\_2020-11-20-12](https://thredds-iow.io-warnemuende.de/thredds/catalogs/projects/integral/catalog_pocNP_V04R25_3nm_agg_time.html?dataset=IOW-THREDDS-Baltic_pocNP_V04R25_3nm_agg_time_2020-11-20-12) (Neumann, 2021). The climatology is available from Bittig et al. (2023): <https://doi.org/10.1594/PANGAEA.961119>.

## 6 Conclusions

In this work, we developed an extrapolation approach that combines two worlds: models, specifically the distribution and connectivity that exist in model data variability; and observations in that they provide constraints of the real-world picture.

The most notable features of the approach are that it does not tend to give extreme, out-of-range values even with few data constraints and that it provides local error estimates, which reflect both underlying variability, e.g. coast–basin gradients, and observational data constraints. We consider of particular merit the fact that the extrapolation scheme adapts its spatial scales to the number of observations in a certain area, leading to a sound representation of less uncertainty where more data are available.

Used together with high-quality surface pCO<sub>2</sub> data from SOCAT, we established a climatology that covers the entire Baltic Sea. Given the present data scarcity in the northern basins of the Baltic Sea, a mean pCO<sub>2</sub> climatology is what is reliably achievable today. It can serve as a seasonal baseline of regional pCO<sub>2</sub> evolution that enables assessment of inter-annual variations, e.g. with respect to timing or amplitude. With sustained and/or enhanced observations in such data-poor regions, a fully monthly resolved pCO<sub>2</sub> dataset will become realistic in the future. To this end, SOOP *Tavastland* and SOOP *Silja Serenade* (crossing between Stockholm and Helsinki) provide a positive outlook for uncertainty reduction. Both SOOP lines were recently adopted by ICOS.

Finally, our extrapolation approach as well as the method to establish a climatology are not limited to pCO<sub>2</sub> or the Baltic Sea. Instead, they are transferable to other areas and parameters where the present work can serve as a template.

## Appendix A: Extrapolation approach

### A1 Notation and mathematical background

To represent a spatial dataset at a given time  $t$  with  $m$  spatial points, we use the data vector  $\mathbf{x}_t$  with size  $m \times 1$ . For a spatial time series of  $n$  times, we use the data matrix  $\mathbf{X}$  with size  $m \times n$ , where the  $n$  columns correspond to  $n$  data vectors  $\mathbf{x}_t$ . Data in  $\mathbf{x}_t$  and  $\mathbf{X}$  are the deviation from the space-dependent temporal mean  $\bar{x}_m$ . The empirical data covariance matrix  $\mathbf{Q}_{\mathbf{X}\mathbf{X}}$  with size  $m \times m$  is given by

$$\mathbf{Q}_{\mathbf{X}\mathbf{X}} = \frac{1}{n-1} \mathbf{X}\mathbf{X}^T. \quad (\text{A1})$$

The singular value decomposition (SVD) or empirical orthogonal function (EOF) decomposition of the data matrix  $\mathbf{X}$  decomposes  $\mathbf{X}$  into a series of orthonormal factors or principal components according to

$$\mathbf{X} = \mathbf{E}\mathbf{\Sigma}\mathbf{A}^T, \quad (\text{A2})$$

where  $\mathbf{E}$  is a  $m \times m$  orthonormal matrix,  $\mathbf{\Sigma}$  a  $m \times n$  rectangular diagonal matrix, and  $\mathbf{A}$  a  $n \times n$  orthonormal matrix. With the

above convention on  $m$  and  $n$ , the column vectors  $\mathbf{e}_i$  (with size  $m \times 1$ ) in  $\mathbf{E}$  represent spatial patterns, while the column vectors  $\mathbf{a}_i$  (with size  $n \times 1$ ) in  $\mathbf{A}$  represent amplitude time series. They are also called left-singular vectors and right-singular vectors, respectively. The diagonal elements of  $\mathbf{\Sigma}$  are the so-called singular values  $s_i$  of  $\mathbf{X}$ . Equation (A2) can be written in vector form as

$$\mathbf{x}_t = \mathbf{E}\mathbf{\Sigma}\mathbf{a}_t = \mathbf{E}\boldsymbol{\alpha}_t, \quad (\text{A3})$$

where  $\boldsymbol{\alpha}_t = \mathbf{\Sigma}\mathbf{a}_t$  is the so-called dimensional amplitude at a given time  $t$ .

Based on a SVD of  $\mathbf{X}$  (Eq. A2), the  $m \times m$  matrix  $\mathbf{B} = \mathbf{X}\mathbf{X}^T$  can be expressed as

$$\mathbf{B} = \mathbf{X}\mathbf{X}^T = \mathbf{E}\mathbf{\Sigma}\mathbf{A}^T\mathbf{A}\mathbf{\Sigma}^T\mathbf{E}^T = \mathbf{E}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{E}^T = \mathbf{E}\tilde{\boldsymbol{\Lambda}}\mathbf{E}^T, \quad (\text{A4})$$

where  $\tilde{\boldsymbol{\Lambda}}$  is a diagonal matrix whose diagonal elements are  $\tilde{\lambda}_i = s_i^2$ . Equation (A4) states the eigendecomposition of  $\mathbf{B}$ . Since matrix  $\mathbf{B}$  is proportional to the covariance matrix  $\mathbf{Q}$  (Eq. A1), we can state the eigendecomposition of the data covariance matrix  $\mathbf{Q}$  as

$$\mathbf{Q} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^T, \quad (\text{A5})$$

with corresponding eigenvalue problem

$$\mathbf{Q}\mathbf{E} = \boldsymbol{\Lambda}\mathbf{E} \quad (\text{A6})$$

and

$$\boldsymbol{\Lambda} = \frac{1}{n-1} \tilde{\boldsymbol{\Lambda}}. \quad (\text{A7})$$

The spatial patterns  $\mathbf{e}_i$  with size  $m \times 1$  are eigenvectors of the data covariance matrix, and the diagonal elements  $\lambda_i$  of  $\boldsymbol{\Lambda}$  the corresponding eigenvalues  $\lambda_i$ , which give the amount of variance associated with each  $\mathbf{e}_i$ . From the set of eigenvectors  $\mathbf{e}_i$  and their corresponding amount of variance  $\lambda_i$ , the spatial distribution  $\mathbf{x}_t$  at a given time  $t$  can be reconstructed by determination of a suitable amplitude vector  $\mathbf{a}_t$  (Eq. A3) (see textbooks of linear algebra or statistical data analysis, e.g. Dommenges, 2015).

### A2 Truncation of eigenvalue modes

For practical purposes, reconstruction often uses only the first  $l$  leading eigenvectors instead of all  $m$  eigenvectors  $\mathbf{e}_i$  ( $1 \leq l \leq m$ ). That is,  $\mathbf{E}$ ,  $\boldsymbol{\Lambda}$ , and  $\mathbf{A}$  are split into a matrix  $\mathbf{E}$ ,  $\boldsymbol{\Lambda}$ , and  $\mathbf{A}$  of size  $m \times l$ ,  $l \times l$ , and  $n \times l$ , respectively, which contain the first  $l$  leading eigenvector modes, and a matrix  $\mathbf{E}'$ ,  $\boldsymbol{\Lambda}'$ , and  $\mathbf{A}'$ , which contain the remaining  $m-l$  eigenvector modes. Equations (A2), (A3), and (A5) thus become

$$\mathbf{X} = \mathbf{E}\mathbf{\Sigma}\mathbf{A}^T + \mathbf{E}'\boldsymbol{\Sigma}'\mathbf{A}'^T, \quad (\text{A8})$$

$$\mathbf{x}_t = \mathbf{E}\mathbf{\Sigma}\mathbf{a}_t + \mathbf{E}'\boldsymbol{\Sigma}'\mathbf{a}'_t, \quad \text{and} \quad (\text{A9})$$

$$\mathbf{Q} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^T + \mathbf{E}'\boldsymbol{\Lambda}'\mathbf{E}'^T = \mathbf{P} + \mathbf{P}'. \quad (\text{A10})$$

The leading eigenvector modes  $\mathbf{e}_i$  cover most of the data variance and often represent more “large-scale” spatial patterns,

while the modes higher than  $l$  contain only a small amount of the data variance and can be seen as “small-scale” spatial patterns (Lorenz, 1956; Kaplan et al., 2000; Dommenges, 2015). Correspondingly, the first and second terms in Eq. (A10) give the large-scale and small-scale variances  $\mathbf{P}$  and  $\mathbf{P}'$ , respectively, whereas in Eq. (A8) and its vector form Eq. (A9) they give the large-scale and small-scale contributions, respectively, to the spatial data vector  $\mathbf{x}_t$ .

This split can also be interpreted as decomposition into a “signal” part and a “noise” part. Equations (A8)–(A10) with only the  $l$  leading modes (i.e. with only the first term) are called a truncated EOF reconstruction in which dimensionality is reduced from  $m$  to  $l$  modes. Higher-order modes (in  $\mathbf{E}'$  and  $\mathbf{\Lambda}'$ , i.e. the second term) are assumed to be dominated by noise and error and are discarded in a truncated eigenvector or EOF reconstruction (Lorenz, 1956; Kaplan et al., 2000).

### A3 Extrapolation from observational data

To represent a set of  $k$  observations at a given time  $t$ , we use the observational data vector  $\mathbf{y}_t$  and the observational error  $\sigma_t$ , both with size  $k \times 1$ . The transfer operator  $\mathbf{H}_t$  is then a “sampling” or observation function that depends on the spatial configuration of the observation points and samples from the data vector  $\mathbf{x}_t$  of size  $m \times 1$  so that the result  $\hat{\mathbf{x}}_t$ ,

$$\hat{\mathbf{x}}_t = \mathbf{H}_t \mathbf{x}_t, \quad (\text{A11})$$

has the same size  $k \times 1$  as the observations. To be comparable with  $\hat{\mathbf{x}}_t$ , the observational data  $\mathbf{y}_t$  are expressed as a deviation from the space-dependent temporal mean  $\bar{\mathbf{x}}_m$  too (Elken et al., 2019).

The eigenvalue reconstruction in truncated form (Eq. A9) then becomes

$$\mathbf{H}_t \mathbf{x}_t = \mathbf{H}_t \mathbf{E} \boldsymbol{\Sigma} \boldsymbol{\alpha}_t. \quad (\text{A12})$$

Minimization of a suitable cost function  $Q$ , e.g. for least-squares optimization as in Elken et al. (2019),

$$Q(\boldsymbol{\alpha}_t) = (\mathbf{H}_t \mathbf{E} \boldsymbol{\alpha}_t - \mathbf{y}_t)^T (\mathbf{H}_t \mathbf{E} \boldsymbol{\alpha}_t - \mathbf{y}_t), \quad (\text{A13})$$

$$\boldsymbol{\alpha}_t = \boldsymbol{\Sigma} \boldsymbol{\alpha}_t, \quad (\text{A14})$$

yields a system of  $l$  linear equations with  $l$  unknowns (Elken et al., 2019):

$$\mathbf{D} \boldsymbol{\alpha}_t = \mathbf{h}, \quad (\text{A15})$$

$$\mathbf{D} = \mathbf{E}^T \mathbf{H}_t^T \mathbf{H}_t \mathbf{E} \quad (\text{A16})$$

$$\mathbf{h} = \mathbf{E}^T \mathbf{H}_t^T \mathbf{y}_t, \quad (\text{A17})$$

which provides the observation-based eigenvector amplitudes  $\boldsymbol{\alpha}_t$  or  $\mathbf{a}_t$  (Eq. A14), respectively.

With Eq. (A9) in truncated form and Eq. (A14), the data vector  $\mathbf{x}_t$  can be obtained, i.e. an extrapolation from  $k \times 1$  observational data points  $\mathbf{y}_t$  to the entire spatial domain of  $m \times 1$  data points  $\mathbf{x}_t$  with the help of the truncated eigenvalue decomposition of  $m \times l$  eigenvectors and  $l$  eigenvalues  $\mathbf{e}_i$  and  $\lambda_i$ , or  $\mathbf{E}$  and  $\mathbf{\Lambda}$ , respectively.

### A4 Error considerations

Both “observational error” and “representational error” impact the determination of the eigenvector amplitudes  $\mathbf{a}_t$ . Observational error  $\sigma_t$  includes both instrumental and sampling error and can be used to limit the impact of observational data constraints  $\mathbf{y}_k$  (Kaplan et al., 2000). Representational error is the error made by truncation, i.e. by using only the  $l$  largest, leading eigenvectors and neglecting the remainder of the spectrum, which can be derived from the “small-scale” covariance  $\mathbf{P}'$  contribution (Kaplan et al., 2000).

The error covariance matrix  $\mathcal{R}$  is the sum of the observational data error covariance  $\mathbf{S}$  and a term that accounts for the covariance created in the truncated modes  $\mathbf{E}'$  not resolved by the analysis:

$$\mathcal{R} = \mathbf{S} + \mathbf{H}_t \mathbf{E}' \boldsymbol{\Lambda}' \mathbf{E}'^T \mathbf{H}_t^T = \mathbf{S} + \mathbf{H}_t \mathbf{P}' \mathbf{H}_t^T, \quad (\text{A18})$$

where  $\mathbf{S}$  is a diagonal matrix with the observations’ variance  $\sigma_t^2$  on the diagonal and  $\mathbf{H}_t$  is the aforementioned sampling operator (Kaplan et al., 2000).

By addition of constraints on the cost function  $Q$  (a) to limit the estimated eigenvector amplitudes based on the amount of variance they explain in the original eigenvalue decomposition and (b) that  $\mathbf{a}_t$  determination is limited by the uncertainty of observations and from truncation (see above), Eq. (A13) is modified to (Kaplan et al., 2000)

$$Q(\boldsymbol{\alpha}_t) = (\mathbf{H}_t \mathbf{E} \boldsymbol{\alpha}_t - \mathbf{y}_t)^T \mathcal{R}^{-1} (\mathbf{H}_t \mathbf{E} \boldsymbol{\alpha}_t - \mathbf{y}_t) + \boldsymbol{\alpha}_t^T \boldsymbol{\Lambda}^{-1} \boldsymbol{\alpha}_t, \quad (\text{A19})$$

and the solution Eqs. (A15)–(A17) thus become (Kaplan et al., 2000)

$$\mathbf{D} \boldsymbol{\alpha}_t = \mathbf{h}, \quad (\text{A20})$$

$$\mathbf{D} = \mathbf{E}^T \mathbf{H}_t^T \mathcal{R}^{-1} \mathbf{H}_t \mathbf{E} + \boldsymbol{\Lambda}^{-1} \quad (\text{A21})$$

$$\mathbf{h} = \mathbf{E}^T \mathbf{H}_t^T \mathcal{R}^{-1} \mathbf{y}_t. \quad (\text{A22})$$

On the other hand, the large-scale portion of the data covariance  $\mathbf{Q}$  (Eq. A10) can be obtained from this solution according to Kaplan et al. (2000):

$$\mathbf{P} = \mathbf{E} \mathbf{D}^{-1} \mathbf{E}^T. \quad (\text{A23})$$

$\mathbf{P}$  can be seen as the mapping uncertainty of the extrapolation.

Note that the calculation of  $\mathbf{P}$  only requires the sampling operator  $\mathbf{H}_t$ , i.e. where samples are present, as well as an observational error estimate  $\mathbf{S}$ , but not the actual observations or amplitudes at these locations. In the absence of any observation,  $\mathbf{H}_t$  contains only zeros and Eqs. (A21) and (A23) then simplify to  $\mathbf{D} = \boldsymbol{\Lambda}^{-1}$  and  $\mathbf{P} = \mathbf{E} \boldsymbol{\Lambda} \mathbf{E}^T$  (compare Eq. A10), respectively. Addition of observational constraints thus reduces the mapping uncertainty  $\mathbf{P}$ .

For our purposes, we assume that *off-diagonal* elements in the covariance  $\mathbf{P}'$  are negligible, i.e. that the truncated eigenvector modes  $\mathbf{E}'$  are sufficiently small-scale and that their

variances  $\Lambda'$  are small enough that they do not show a noticeable covariance contribution (Eq. A18). We therefore approximate  $\mathbf{P}'$  by a diagonal matrix whose diagonal elements are obtained from rearrangement of Eq. (A10):

$$\text{diag}(\mathbf{P}') = \text{diag}(\mathbf{Q}) - \text{diag}(\mathbf{E}\mathbf{\Lambda}\mathbf{E}^T), \quad (\text{A24})$$

where  $\mathbf{Q}$  is accessible from the data  $\mathbf{X}$  (Eq. A1) and the second term from the truncated eigenvalue decomposition.

### A5 Number of significant modes

A critical aspect before any extrapolation from observations is how many modes  $e_i$  to use, or – in other words – where to truncate reconstruction, i.e. how to distinguish between modes representing desired variability and modes representing “noise”. Often, an arbitrary threshold of the total variance covered by the  $l$  leading modes is chosen (e.g. Kaplan et al., 2000).

Here, we apply the DINEOF (Data Interpolating Empirical Orthogonal Functions) variant of SVD or EOF decomposition of the dataset  $\mathbf{X}$  to find the  $l_{\text{max}}$  leading eigenvectors  $e_i$ . The number  $l_{\text{max}}$  of significant modes is determined by a cross-validation procedure from the data  $\mathbf{X}$  (Beckers and Rixen, 2003), and we use this  $l_{\text{max}}$  as the upper bound for  $l$  in our ensemble approach (see below).

### A6 Ensemble approach for robustness and locality

Depending on the spatial distribution or clustering of observations  $y_t$ , not all amplitudes of the  $l_{\text{max}}$  spatial eigenvectors  $e_i$  may be well constrained. In some configurations, e.g. with few or clustered observations, where  $\mathbf{H}_t$  selects only a few of the  $m$  spatial data points, a reconstruction with fewer modes than  $l_{\text{max}}$  may give a solution with a smaller mapping uncertainty  $\mathbf{P}$  than the truncated reconstruction with all  $l_{\text{max}}$  modes  $e_i$ .

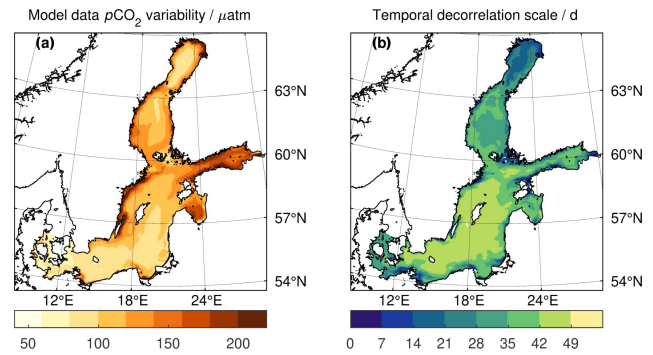
We therefore use an ensemble approach over the truncated reconstructions, where we vary  $l$  from just 1 mode to  $l_{\text{max}}$  modes:

$$\bar{\mathcal{X}} = \frac{\sum_{l=1}^{l_{\text{max}}} \mathbf{w}_l \cdot \mathcal{X}_l}{\sum_{l=1}^{l_{\text{max}}} \mathbf{w}_l}, \quad (\text{A25})$$

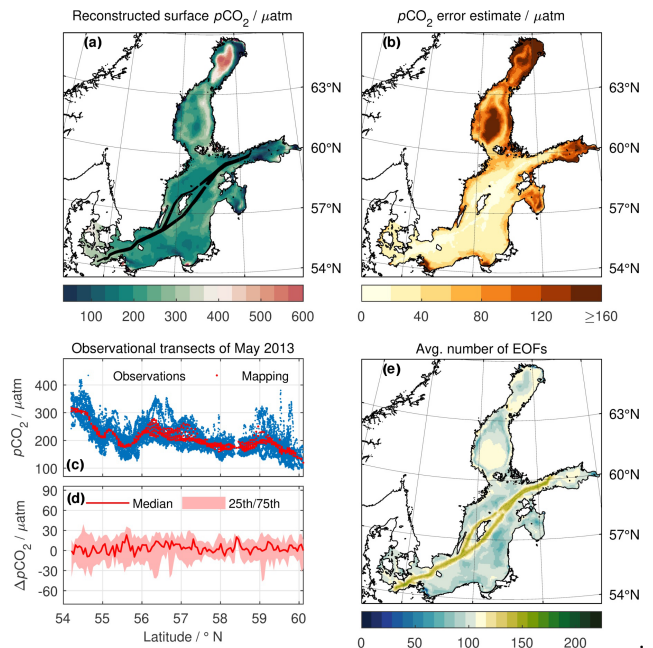
where the weights  $\mathbf{w}_l$  (with size  $m \times 1$ ) are based on the diagonal elements of  $\mathbf{P}$ , i.e. the variance  $\sigma_{P,t}^2$  or mapping uncertainty of the extrapolation:

$$\mathbf{w}_l = 1 \left/ \frac{\sigma_{P,t}^2}{\sum_{l=1}^m \sigma_{P,t}^2} \right. . \quad (\text{A26})$$

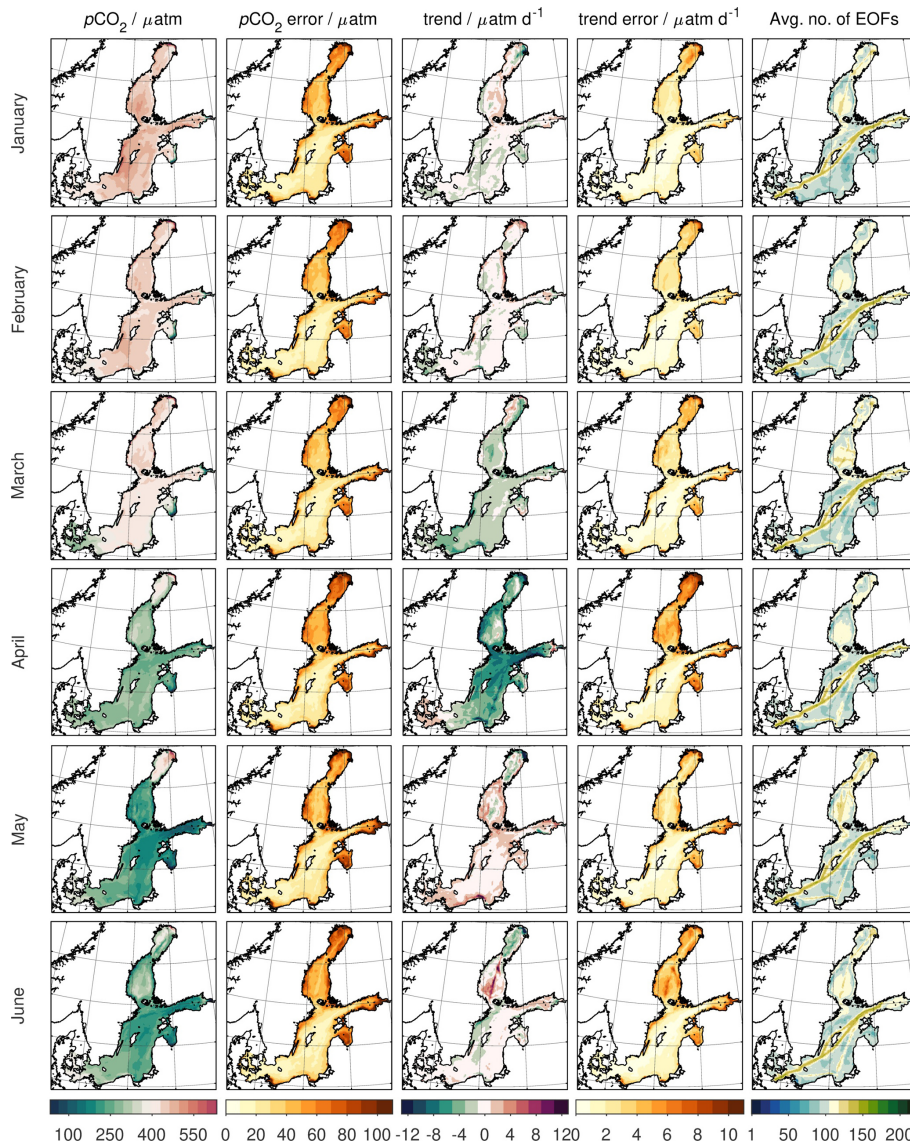
As the mapped variance  $\text{diag}(\mathbf{P})$  varies in magnitude with the number of eigenvectors included, the weights  $\mathbf{w}_l$  for a given number of modes  $l$  are normalized by the total variance over all  $m$  data points, which gives a normalized spatial weight



**Figure A1.** Model data  $\mathbf{X}$  characteristics: (a)  $p\text{CO}_2$  variability shown by the square root of the model data covariance matrix  $\mathbf{Q}$ 's diagonal elements, i.e. the standard deviation of the  $p\text{CO}_2$  data  $\mathbf{X}$  at each location. (b)  $p\text{CO}_2$  variability shown by the temporal decorrelation timescale, i.e. the time (d) at which the lagged autocorrelation drops below a threshold of 0.63 correlation. Less than 1 % of the locations have a temporal decorrelation scale of  $\leq 7$  d, while about 25 % are within  $\leq 30$  d. To adequately reflect the  $p\text{CO}_2$  data dynamics in the variability pattern extraction (Sect. 2.2.1), a weekly aggregation of the model data  $\mathbf{X}$  was therefore chosen compared to e.g. daily or monthly model output.



**Figure A2.** Same as Fig. 3 but without temporally collating the observations to the middle of the month (details in Sect. A7). The general picture of (a) reconstructed  $p\text{CO}_2$   $x_{t,\text{reconstr}}$ , (b)  $p\text{CO}_2$  error estimate  $\sigma_{t,\text{reconstr}}$ , and (e) average number of patterns  $\bar{l}$  of the reconstruction is comparable to the temporally coherent reconstruction (Fig. 3a, b, g, respectively), but (c) mapped  $p\text{CO}_2$  (red) gives only one uniform value throughout the month, and (d) differences to observations are therefore increased (compare Fig. 3e, f).



**Figure A3.** Monthly climatology  $\mathbf{Y}$  of the surface  $p\text{CO}_2$ ,  $p\text{CO}_2$  error estimate,  $p\text{CO}_2$  trend,  $p\text{CO}_2$  trend error estimate, and average number of EOF patterns in the ensemble (left to right) for the months January–June (top to bottom).

vector  $\mathbf{w}_l$  (Eq. A26). For a given spatial data point  $m$ , preference is thus given to reconstructions where it is among the better-constrained ones for the given number of modes  $l$ .

For  $\mathcal{X}_l$ , each using only the first  $l$  ( $\leq l_{\max}$ ) eigenvector modes (Eq. A25), we insert the following.

- The reconstructed data vector  $\mathbf{x}_t$  (from Eq. A9 in truncated form) to yield as  $\overline{\mathcal{X}}$  the ensemble mean  $\overline{\mathbf{x}}_t$
- The term  $(\overline{\mathbf{x}}_t - \mathbf{x}_t)^2$  to yield as  $\overline{\mathcal{X}}$  the biased weighted sample variance about the ensemble mean  $\sigma_{t,\text{mean}}^2$
- The mapping variance  $\sigma_{P',t}^2$  (i.e. the diagonal of  $\mathbf{P}'$ ; Eq. A23) to yield as  $\overline{\mathcal{X}}$  the remaining mean “large-scale” variance  $\sigma_{P',t}^2$

- The approximation of the truncated variance  $\sigma_{P',t}^2$  (i.e. the diagonal of  $\mathbf{P}'$ ; Eq. A24) to yield as  $\overline{\mathcal{X}}$  the mean “small-scale” variance  $\sigma_{P',t}^2$
- The number of modes  $l$  itself to yield as  $\overline{\mathcal{X}}$  the average number of modes  $\overline{l}$  used in the reconstruction as a spatial vector of size  $m \times 1$  each, i.e. on the full spatial domain

The reconstructed data vector  $\mathbf{x}_{t,\text{reconstr}}$  and the total variance of reconstruction  $\sigma_{t,\text{reconstr}}^2$  are given by

$$\mathbf{x}_{t,\text{reconstr}} = \overline{\mathbf{x}}_t, \quad (\text{A27})$$

$$\sigma_{t,\text{reconstr}}^2 = \sigma_{t,\text{mean}}^2 + \overline{\sigma_{P',t}^2} + \overline{\sigma_{P',t}^2}. \quad (\text{A28})$$



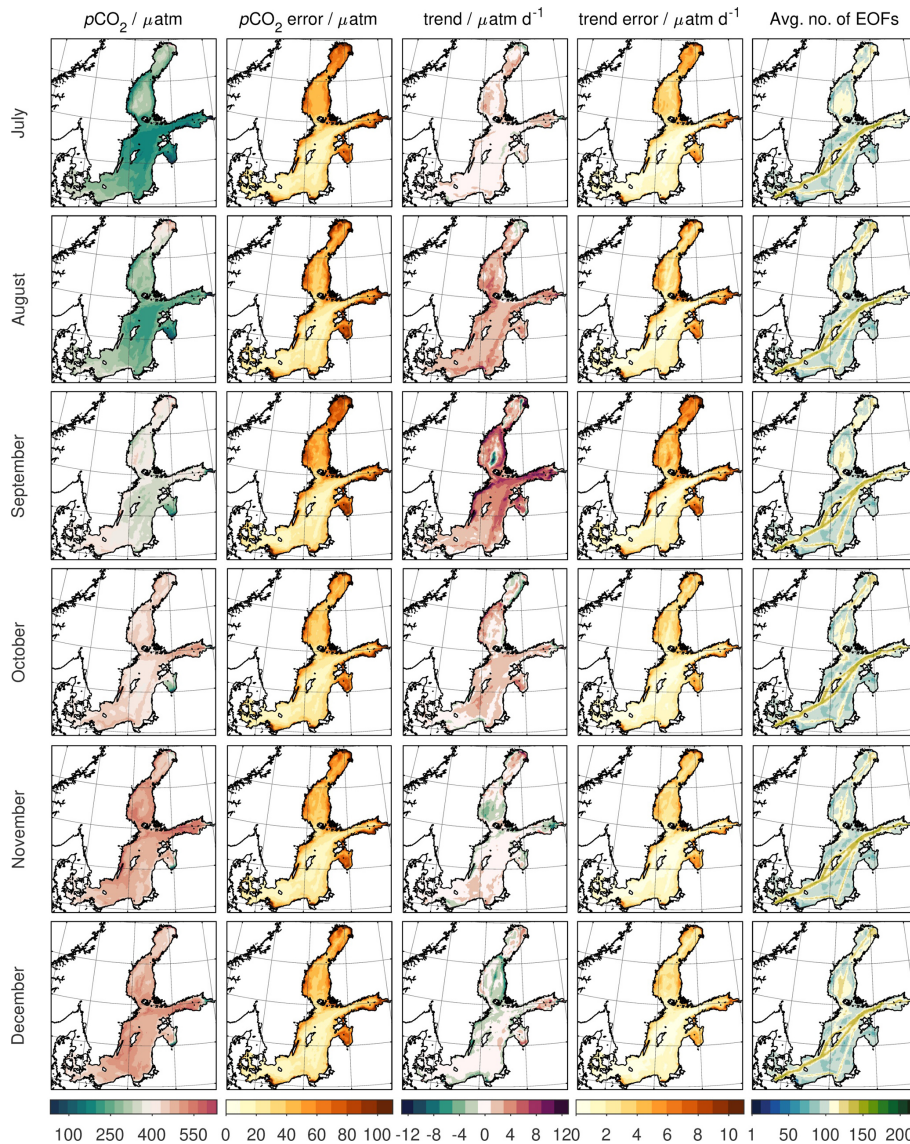


Figure A4. Same as Fig. A3 but for July–December.

### A7 Temporal coherence

So far, we have considered the observations  $y_t$  and the reconstructed data vector  $x_t$  to originate and be valid for a given time  $t$ , respectively. However, in practice, observations  $y$  will spread over a certain time extent  $\Delta t$ , e.g. days for a single SOOP transect to weeks for a basin-covering cruise. If this time extent (e.g. duration of a cruise) is comparable to the timescale of the system’s dynamics (e.g. spring surface warming or bloom onset), distortions in the reconstructed fields  $x_t$  may occur (Elken et al., 2019) as different “states” of the system could be sampled at different space–time locations.

To collate temporally extended observations into a common synoptic reconstruction without artefacts, Elken et al. (2019) added a linear temporal (short-term) trend for each

eigenvector mode to the optimization, thus collating observations made at different observation times  $t$  to a common reference time instance  $t_0$ .

To this end,

- the time difference  $\Delta t$  between observation and reference times is introduced:

$$\Delta t = t - t_0. \tag{A29}$$

- The sampling operator  $\mathbf{H}_t$  is extended to not only sample from the eigenvector  $e_i$  the value  $\chi$  corresponding to the observation, but also the value times the time difference,  $\chi \cdot \Delta t$ , corresponding to a (short-term) time gradient.  $\mathbf{H}_t \mathbf{E}$  thus becomes a matrix of size  $k \times 2l$ .

- The  $l$  eigenvalues are replicated with a constant scaling factor, so that  $\mathbf{A}$  becomes a  $2l \times 2l$  diagonal matrix. That is, both the value (first  $l$  eigenvalues) and its time difference (second  $l$  eigenvalues times the scaling factor) follow the same eigenvector  $\mathbf{e}_i$  order and importance, with the scaling factor to determine the balance between both. Based on preliminary tests, we used an empirical scaling between the spatial value pattern and temporal trend pattern of  $9 \times 10^{-4}$ , equal to a ratio of approximately  $1 \mu\text{atm}$  per 33 d.
- The eigenvalue amplitude vector  $\mathbf{a}_t$  is doubled in size to  $2l \times 1$ , where the first  $l$  values continue to represent the eigenvectors'  $\mathbf{e}_i$  amplitudes, while the second  $l$  values represent their temporal derivative.

Thus, the system of linear equations (Eqs. A20–A22) becomes a system of  $2l$  equations with  $2l$  unknowns, and the reconstruction (Eq. A9) provides both an extrapolation to the entire spatial domain of  $m \times 1$  data locations at reference time  $t_o$  as well as a temporal trend of change in the  $m \times 1$  data locations at reference time  $t_o$  (Elken et al., 2019).

## Appendix B: Seasonal $p\text{CO}_2$ from value and linear trend climatology by cubic Hermite spline calculation

With a  $p\text{CO}_2$  value  $x$  and linear  $p\text{CO}_2$  trend  $d$  given for each month of the climatology, the seasonal evolution of  $p\text{CO}_2$  at a given location  $m$  can be calculated as a cubic Hermite spline according to

$$x_t = h_{00}(\gamma) \cdot x_k + h_{10}(\gamma) \cdot d_k \cdot (t_{k+1} - t_k) + \dots \\ h_{01}(\gamma) \cdot x_{k+1} + h_{11}(\gamma) \cdot d_{k+1} \cdot (t_{k+1} - t_k), \quad (\text{B1})$$

with

$$h_{00}(\gamma) = 2\gamma^3 - 3\gamma^2 + 1, \quad (\text{B2})$$

$$h_{10}(\gamma) = \gamma^3 - 2\gamma^2 + \gamma, \quad (\text{B3})$$

$$h_{01}(\gamma) = -2\gamma^3 + 3\gamma^2, \quad (\text{B4})$$

$$h_{11}(\gamma) = \gamma^3 - \gamma^2, \text{ and} \quad (\text{B5})$$

$$\gamma = \frac{t - t_k}{t_{k+1} - t_k}. \quad (\text{B6})$$

Here,  $t_k$  and  $t_{k+1}$  are the times of the climatological month before and after the time  $t$  of interest, and  $\gamma$  gives the normalized time.  $h(\gamma)$  are Hermite basis functions and  $x_k$  and  $x_{k+1}$  as well as  $d_k$  and  $d_{k+1}$  are the associated monthly  $p\text{CO}_2$  value and  $p\text{CO}_2$  trend, respectively, which determine the seasonal  $p\text{CO}_2$   $x_t$  at time  $t$ .

**Author contributions.** HCB and GR conceived the study. The method was developed by HCB with important input by EJ and GR. TN performed the model simulations and HCB the analysis. HCB led the manuscript writing with contributions by all the co-authors.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

**Acknowledgements.** Efforts by Anna Willstrand-Wranne (SMHI) and Tobias Steinhoff (GEOMAR/NORCE) around SOOP *Tavastland* observations are highly acknowledged. The Surface Ocean  $\text{CO}_2$  Atlas (SOCAT) is an international effort endorsed by the International Ocean Carbon Coordination Project (IOCCP), the Surface Ocean Lower Atmosphere Study (SOLAS), and the Integrated Marine Biosphere Research (IMBeR) programme to deliver a uniformly quality-controlled surface ocean  $\text{CO}_2$  database. The many researchers and funding agencies responsible for the collection of data and quality control are thanked for their contributions to SOCAT. The authors want to thank two anonymous reviewers and the handling editor for their valuable and constructive comments which improved the manuscript.

**Financial support.** This work was funded by the projects C-SCOPE (grant no. 03F0877D) and BONUS INTEGRAL (grant no. 03F0773A), which received funding from BONUS (Art. 185), funded jointly by the EU, the German Federal Ministry of Education and Research, the Swedish Research Council Formas, the Academy of Finland, the Polish National Centre for Research and Development, and the Estonian Research Council. Computational power was provided by the North German Supercomputing Alliance (HLRN). Measurements on SOOP *Finnmaid* were temporarily (2009–2011) funded by the German Federal Ministry of Education and Research in the framework of BONUS projects Baltic-C (grant no. 03F0486A), Baltic Gas (grant no. 03F0488B), and, since 2012, ICOS-D (grant nos. 01LK1101F and 01LK1224D). Additional support came from the JERICO-S3 project funded by the European Commission's H2020 Framework Programme under grant agreement no. 871153.

The publication of this article was funded by the Open Access Fund of the Leibniz Association.

**Review statement.** This paper was edited by Giuseppe M. R. Manzella and reviewed by two anonymous referees.

## References

- Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A.,

- Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibánhez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., and Xu, S.: A multi-decade record of high-quality *f*CO<sub>2</sub> data in version 3 of the Surface Ocean CO<sub>2</sub> Atlas (SOCAT), *Earth Syst. Sci. Data*, 8, 383–413, <https://doi.org/10.5194/essd-8-383-2016>, 2016.
- Bakker, D. C. E., Alin, S. R., Becker, M., Bittig, H. C., Castaño-Primo, R., Feely, R. A., Gkritzalis, T., Kadono, K., Kozyr, A., Lauvset, S. K., Metzl, N., Munro, D. R., Nakaoka, S., Nojiri, Y., O'Brien, K. M., Olsen, A., Pfeil, B., Pierrot, D., Steinhoff, T., Sullivan, K. F., Sutton, A. J., Sweeney, C., Tilbrook, B., Wada, C., Wanninkhof, R., Willstrand Wranne, A., Akl, J., Apelthun, L. B., Bates, N. R., Beatty, C. M., Burger, E. F., Cai, W.-J., Cosca, C. E., Corredor, J. E., Cronin, M., Cross, J. N., De Carlo, E. H., DeGrandpre, M. D., Emerson, S. R., Enright, M. P., Enyo, K., Evans, W., Frangoulis, C., Fransson, A., García-Ibáñez, M. I., Gehrung, M., Giannoudi, L., Glockzin, M., Hales, B., Howden, S. D., Hunt, C. W., Ibánhez, J. S. P., Jones, S. D., Kamb, L., Körtzinger, A., Landa, C. S., Landschützer, P., Lefèvre, N., Lo Monaco, C., Macovei, V. A., Maenner Jones, S., Meinig, C., Millero, F. J., Monacci, N. M., Mordy, C., Morell, J. M., Murata, A., Musielewicz, S., Neill, C., Newberger, T., Nomura, D., Ohman, M., Ono, T., Passmore, A., Petersen, W., Petihakis, G., Perivoliotis, L., Plueddemann, A. J., Rehder, G., Reynaud, T., Rodriguez, C., Ross, A. C., Rutgersson, A., Sabine, C. L., Salisbury, J. E., Schlitzer, R., Send, U., Skjelvan, I., Stamataki, N., Sutherland, S. C., Sweeney, C., Tadokoro, K., Tanhua, T., Telszewski, M., Trull, T., Vandemark, D., van Ooijen, E., Vaynsnova, Y. G., Wang, H., Weller, R. A., Whitehead, C., and Wilson, D.: Surface Ocean CO<sub>2</sub> Atlas Database Version 2022 (SOCATv2022) (NCEI Accession 0253659), NOAA National Centers for Environmental Information [data set], <https://doi.org/10.25921/1h9f-nb73>, 2022.
- Becker, M., Olsen, A., Landschützer, P., Omar, A., Rehder, G., Rödenbeck, C., and Skjelvan, I.: The northern European shelf as an increasing net sink for CO<sub>2</sub>, *Biogeosciences*, 18, 1127–1147, <https://doi.org/10.5194/bg-18-1127-2021>, 2021.
- Beckers, J. M. and Rixen, M.: EOF Calculations and Data Filling from Incomplete Oceanographic Datasets, *J. Atmos. Ocean. Tech.*, 20, 1839–1856, [https://doi.org/10.1175/1520-0426\(2003\)020<1839:ECADFF>2.0.CO;2](https://doi.org/10.1175/1520-0426(2003)020<1839:ECADFF>2.0.CO;2), 2003.
- Bittig, H. C., Jacobs, E., Neumann, T., and Rehder, G.: A regional pCO<sub>2</sub> climatology of the Baltic Sea, PANGAEA [data set], <https://doi.org/10.1594/PANGAEA.961119>, 2023.
- Carstensen, J. and Duarte, C. M.: Drivers of pH Variability in Coastal Ecosystems, *Environ. Sci. Technol.*, 53, 4020–4029, <https://doi.org/10.1021/acs.est.8b03655>, 2019.
- Davis, R. E.: Predictability of Sea Surface Temperature and Sea Level Pressure Anomalies over the North Pacific Ocean, *J. Phys. Oceanogr.*, 6, 249–266, [https://doi.org/10.1175/1520-0485\(1976\)006<0249:POSSTA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0249:POSSTA>2.0.CO;2), 1976.
- DeVries, T.: The Ocean Carbon Cycle, *Annu. Rev. Environ. Resour.*, 47, 317–341, <https://doi.org/10.1146/annurev-environ-120920-111307>, 2022.
- Dommenget, D.: An Introduction to Statistical Analysis in Climate Research, <https://users.monash.edu.au/~dietmard/teaching/honours.statistics/dommenget.statistics.lecture.notes.pdf> (last access: 31 January 2024), 2015.
- Eilola, K., Gustafsson, B. G., Kuznetsov, I., Meier, H. E. M., Neumann, T., and Savchuk, O. P.: Evaluation of biogeochemical cycles in an ensemble of three state-of-the-art numerical models of the Baltic Sea, *J. Marine Syst.*, 88, 267–284, <https://doi.org/10.1016/j.jmarsys.2011.05.004>, 2011.
- Elken, J., Zujev, M., She, J., and Lagema, P.: Reconstruction of Large-Scale Sea Surface Temperature and Salinity Fields Using Sub-Regional EOF Patterns From Models, *Front. Earth Sci.*, 7, 232, <https://doi.org/10.3389/feart.2019.00232>, 2019.
- Friedlingstein, P., O'Sullivan, M., Jones, M. W., Andrew, R. M., Gregor, L., Hauck, J., Le Quéré, C., Luijckx, I. T., Olsen, A., Peters, G. P., Peters, W., Pongratz, J., Schwingshackl, C., Sitch, S., Canadell, J. G., Ciais, P., Jackson, R. B., Alin, S. R., Alkama, R., Arneeth, A., Arora, V. K., Bates, N. R., Becker, M., Bellouin, N., Bittig, H. C., Bopp, L., Chevallier, F., Chini, L. P., Cronin, M., Evans, W., Falk, S., Feely, R. A., Gasser, T., Gehlen, M., Gkritzalis, T., Gloege, L., Grassi, G., Gruber, N., Gürses, Ö., Harris, I., Hefner, M., Houghton, R. A., Hurtt, G. C., Iida, Y., Ilyina, T., Jain, A. K., Jersild, A., Kadono, K., Kato, E., Kennedy, D., Klein Goldewijk, K., Knauer, J., Korsbakken, J. I., Landschützer, P., Lefèvre, N., Lindsay, K., Liu, J., Liu, Z., Marland, G., Mayot, N., McGrath, M. J., Metzl, N., Monacci, N. M., Munro, D. R., Nakaoka, S.-I., Niwa, Y., O'Brien, K., Ono, T., Palmer, P. I., Pan, N., Pierrot, D., Pockock, K., Poulter, B., Resplandy, L., Robertson, E., Rödenbeck, C., Rodriguez, C., Rosan, T. M., Schwinger, J., Séférian, R., Shutler, J. D., Skjelvan, I., Steinhoff, T., Sun, Q., Sutton, A. J., Sweeney, C., Takao, S., Tanhua, T., Tans, P. P., Tian, X., Tian, H., Tilbrook, B., Tsujino, H., Tubiello, F., van der Werf, G. R., Walker, A. P., Wanninkhof, R., Whitehead, C., Willstrand Wranne, A., Wright, R., Yuan, W., Yue, C., Yue, X., Zaehle, S., Zeng, J., and Zheng, B.: Global Carbon Budget 2022, *Earth Syst. Sci. Data*, 14, 4811–4900, <https://doi.org/10.5194/essd-14-4811-2022>, 2022.
- GESAMP: High Level Review of a Wide Range of Proposed Marine Geoengineering Techniques, Rep. Stud. GESAMP No. 98, (IMO/FAO/UNESCO-IOC/UNIDO/WMO/IAEA/UN/UN Environment/UNDP/ISA Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection), <http://www.gesamp.org/publications/high-level-review-of-a-wide-range-of-proposed-marine-geoengineering-techniques> (last access: 31 January 2024), 2019.
- Gülzow, W., Rehder, G., Schneider, B., Schneider v. Deimling, J., and Sadkowiak, B.: A new method for continuous measurement of methane and carbon dioxide in surface waters using off-axis integrated cavity output spectroscopy (ICOS): An example from the Baltic Sea, *Limnol. Oceanogr.-Methods*, 9, 176–184, <https://doi.org/10.4319/lom.2011.9.176>, 2011.

- Hannachi, A., Jolliffe, I. T., and Stephenson, D. B.: Empirical orthogonal functions and related techniques in atmospheric science: A review, *Int. J. Climatol.*, 27, 1119–1152, <https://doi.org/10.1002/joc.1499>, 2007.
- HELCOM Secretariat: HELCOM subbasins 2017 (level 2), <https://metadata.helcom.fi/geonetwork/srv/eng/catalog.search#/metadata/d4b6296c-fd19-462c-94d2-4c81b9313d77> (last access: 14 May 2019), 2017.
- Humborg, C., Geibel, M. C., Sun, X., McCrackin, M., Morth, C.-M., Stranne, C., Jakobsson, M., Gustafsson, B., Sokolov, A., Norkko, A., and Norkko, J.: High Emissions of Carbon Dioxide and Methane From the Coastal Baltic Sea at the End of a Summer Heat Wave, *Front. Mar. Sci.*, 6, 493, <https://doi.org/10.3389/fmars.2019.00493>, 2019.
- Jacobs, E., Bittig, H. C., Gräwe, U., Graves, C. A., Glockzin, M., Müller, J. D., Schneider, B., and Rehder, G.: Upwelling-induced trace gas dynamics in the Baltic Sea inferred from 8 years of autonomous measurements on a ship of opportunity, *Biogeosciences*, 18, 2679–2709, <https://doi.org/10.5194/bg-18-2679-2021>, 2021.
- Jolliffe, I. T. and Cadima, J.: Principal component analysis: a review and recent developments, *Phil. Trans. R. Soc. A*, 374, 20150202, <https://doi.org/10.1098/rsta.2015.0202>, 2016.
- Kaplan, A., Kushnir, Y., and Cane, M. A.: Reduced Space Optimal Interpolation of Historical Marine Sea Level Pressure: 1854–1992, *J. Climate*, 13, 2987–3002, [https://doi.org/10.1175/1520-0442\(2000\)013<2987:RSOIOH>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<2987:RSOIOH>2.0.CO;2), 2000.
- Kuliński, K., Rehder, G., Asmala, E., Bartosova, A., Carstensen, J., Gustafsson, B., Hall, P. O. J., Humborg, C., Jilbert, T., Jürgens, K., Meier, H. E. M., Müller-Karulis, B., Naumann, M., Olesen, J. E., Savchuk, O., Schramm, A., Slomp, C. P., Sofiev, M., Sobek, A., Szymczycha, B., and Undeman, E.: Biogeochemical functioning of the Baltic Sea, *Earth Syst. Dynam.*, 13, 633–685, <https://doi.org/10.5194/esd-13-633-2022>, 2022.
- Kuznetsov, I. and Neumann, T.: Simulation of carbon dynamics in the Baltic Sea with a 3D model, *J. Marine Syst.*, 111–112, 167–174, <https://doi.org/10.1016/j.jmarsys.2012.10.011>, 2013.
- Körtzinger, A., Thomas, H., Schneider, B., Gronau, N., Mintrop, L., and Duinker, J. C.: At-sea intercomparison of two newly designed underway  $p\text{CO}_2$  systems – Encouraging results, *Mar. Chem.*, 52, 133–145, [https://doi.org/10.1016/0304-4203\(95\)00083-6](https://doi.org/10.1016/0304-4203(95)00083-6), 1996.
- Lan, X., Tans, P., and Thoning, K. W.: Trends in globally-averaged  $\text{CO}_2$  determined from NOAA Global Monitoring Laboratory measurements. Version 2022-11, <https://gml.noaa.gov/ccgg/trends/> (last access: 15 November 2022), 2022.
- Landschützer, P., Gruber, N., Bakker, D. C. E., Schuster, U., Nakaoka, S., Payne, M. R., Sasse, T. P., and Zeng, J.: A neural network-based estimate of the seasonal to inter-annual variability of the Atlantic Ocean carbon sink, *Biogeosciences*, 10, 7793–7815, <https://doi.org/10.5194/bg-10-7793-2013>, 2013.
- Laruelle, G. G., Cai, W.-J., Hu, X., Gruber, N., Mackenzie, F. T., and Regnier, P.: Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide, *Nat. Commun.*, 9, 454, <https://doi.org/10.1038/s41467-017-02738-z>, 2018.
- Lorenz, E. N.: Empirical orthogonal functions and statistical weather prediction, Scientific Report No. 1, Statistical Forecasting Project, Massachusetts Institute of Technology, Department of Meteorology, Cambridge, MT, USA, [https://eapsweb.mit.edu/sites/default/files/Empirical\\_Orthogonal\\_Functions\\_1956.pdf](https://eapsweb.mit.edu/sites/default/files/Empirical_Orthogonal_Functions_1956.pdf) (last access: 31 January 2024), oCLC: 2293210, 1956.
- Luisetti, T., Ferrini, S., Grilli, G., Jickells, T. D., Kennedy, H., Kröger, S., Lorenzoni, I., Milligan, B., van der Molen, J., Parker, R., Pryce, T., Turner, R. K., and Tyllianakis, E.: Climate action requires new accounting guidance and governance frameworks to manage carbon in shelf seas, *Nat. Commun.*, 11, 4599, <https://doi.org/10.1038/s41467-020-18242-w>, 2020.
- Monahan, A. H., Fyfe, J. C., Ambaum, M. H. P., Stephenson, D. B., and North, G. R.: Empirical Orthogonal Functions: The Medium is the Message, *J. Climate*, 22, 6501–6514, <https://doi.org/10.1175/2009JCLI3062.1>, 2009.
- Müller, J. D., Schneider, B., and Rehder, G.: Long-term alkalinity trends in the Baltic Sea and their implications for  $\text{CO}_2$ -induced acidification, *Limnol. Oceanogr.*, 61, 1984–2002, <https://doi.org/10.1002/lno.10349>, 2016.
- Neumann, T.: ERGOM 1.2 model hindcast 1948–2020, IOW [data set], [https://thredds-iow.io-warnemuende.de/thredds/catalogs/projects/integral/catalog\\_pocNP\\_V04R25\\_3nm\\_agg\\_time.html?dataset=IOW-THREDDSS-Baltic\\_pocNP\\_V04R25\\_3nm\\_agg\\_time\\_2020-11-20-12](https://thredds-iow.io-warnemuende.de/thredds/catalogs/projects/integral/catalog_pocNP_V04R25_3nm_agg_time.html?dataset=IOW-THREDDSS-Baltic_pocNP_V04R25_3nm_agg_time_2020-11-20-12) (last access: 31 January 2024), 2021.
- Neumann, T., Siegel, H., and Gerth, M.: A new radiation model for Baltic Sea ecosystem modelling, *J. Marine Syst.*, 152, 83–91, <https://doi.org/10.1016/j.jmarsys.2015.08.001>, 2015.
- Neumann, T., Radtke, H., Cahill, B., Schmidt, M., and Rehder, G.: Non-Redfieldian carbon model for the Baltic Sea (ERGOM version 1.2) – implementation and budget estimates, *Geosci. Model Dev.*, 15, 8473–8540, <https://doi.org/10.5194/gmd-15-8473-2022>, 2022.
- Omstedt, A., Gustafsson, E., and Wesslander, K.: Modelling the uptake and release of carbon dioxide in the Baltic Sea surface water, *Cont. Shelf Res.*, 29, 870–885, <https://doi.org/10.1016/j.csr.2009.01.006>, 2009.
- Parard, G., Charantonis, A. A., and Rutgersson, A.: Using satellite data to estimate partial pressure of  $\text{CO}_2$  in the Baltic Sea, *J. Geophys. Res.-Biogeo.*, 121, 1002–1015, <https://doi.org/10.1002/2015JG003064>, 2016.
- Parard, G., Rutgersson, A., Raj Parampil, S., and Charantonis, A. A.: The potential of using remote sensing data to estimate air–sea  $\text{CO}_2$  exchange in the Baltic Sea, *Earth Syst. Dynam.*, 8, 1093–1106, <https://doi.org/10.5194/esd-8-1093-2017>, 2017.
- Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke, A., Metzl, N., Sabine, C. L., Akl, J., Alin, S. R., Bates, N., Bellerby, R. G. J., Borges, A., Boutin, J., Brown, P. J., Cai, W.-J., Chavez, F. P., Chen, A., Cosca, C., Fassbender, A. J., Feely, R. A., González-Dávila, M., Goyet, C., Hales, B., Hardman-Mountford, N., Heinze, C., Hood, M., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Jones, S. D., Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A., Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A., Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G.-H., Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Santana-Casiano, J. M., Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R., Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas, H., Tilbrook, B., Tjiputra, J., Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J., Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: A uniform, quality controlled

- Surface Ocean CO<sub>2</sub> Atlas (SOCAT), *Earth Syst. Sci. Data*, 5, 125–143, <https://doi.org/10.5194/essd-5-125-2013>, 2013.
- Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T., Olsen, A., Feely, R. A., and Cosca, C. E.: Recommendations for autonomous underway  $p\text{CO}_2$  measuring systems and data-reduction routines, *Deep-Sea Res. Pt. II*, 56, 512–522, <https://doi.org/10.1016/j.dsr2.2008.12.005>, 2009.
- Preisendorfer, R. W.: Principal component analysis in meteorology and oceanography, no. 17 in *Developments in atmospheric science*, Elsevier, Amsterdam, 1988.
- Rehder, G., Velasco-Sobeck, S., Kreuzer, L., Gutierrez Loza, L., Kozirowska, K., Makuch, P., Jacobs, E., Otto, S., Bittig, H., Lainela, S., Honkanen, M., and Kolbe, M.: Continuous nutrient, trace gas and CO<sub>2</sub> measurements during EMB214, May–June 2019, PANGAEA [data set], <https://doi.org/10.1594/PANGAEA.931313>, 2021.
- Rödenbeck, C., Bakker, D. C. E., Gruber, N., Iida, Y., Jacobson, A. R., Jones, S., Landschützer, P., Metzl, N., Nakaoka, S., Olsen, A., Park, G.-H., Peylin, P., Rodgers, K. B., Sasse, T. P., Schuster, U., Shutler, J. D., Valsala, V., Wanninkhof, R., and Zeng, J.: Data-based estimates of the ocean carbon sink variability – first results of the Surface Ocean  $p\text{CO}_2$  Mapping intercomparison (SOCOM), *Biogeosciences*, 12, 7251–7278, <https://doi.org/10.5194/bg-12-7251-2015>, 2015.
- Schneider, B. and Müller, J. D.: *Biogeochemical Transformations in the Baltic Sea: Observations Through Carbon Dioxide Glasses*, Springer Oceanography, Springer International Publishing, <https://doi.org/10.1007/978-3-319-61699-5>, 2018.
- Schneider, B., Kaitala, S., and Maunula, P.: Identification and quantification of plankton bloom events in the Baltic Sea by continuous  $p\text{CO}_2$  and chlorophyll *a* measurements on a cargo ship, *J. Marine Syst.*, 59, 238–248, <https://doi.org/10.1016/j.jmarsys.2005.11.003>, 2006.
- Thomas, H., Bozec, Y., Elkalay, K., and de Baar, H. J. W.: Enhanced Open Ocean Storage of CO<sub>2</sub> from Shelf Sea Pumping, *Science*, 304, 1005–1008, <https://doi.org/10.1126/science.1095491>, 2004.
- Weare, B. C. and Newell, R. E.: Empirical orthogonal analysis of atlantic ocean surface temperatures, *Q. J. Roy. Meteor. Soc.*, 103, 467–478, <https://doi.org/10.1002/qj.49710343707>, 1977.
- Weare, B. C., Navato, A. R., and Newell, R. E.: Empirical Orthogonal Analysis of Pacific Sea Surface Temperatures, *J. Phys. Oceanogr.*, 6, 671–678, [https://doi.org/10.1175/1520-0485\(1976\)006<0671:EOAOPS>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0671:EOAOPS>2.0.CO;2), 1976.
- Wesslander, K., Omstedt, A., and Schneider, B.: Inter-annual and seasonal variations in the air–sea CO<sub>2</sub> balance in the central Baltic Sea and the Kattegat, *Cont. Shelf Res.*, 30, 1511–1521, <https://doi.org/10.1016/j.csr.2010.05.014>, 2010.