https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Equation discovery for climate impact: emulating impact models for unexplored climate scenario with interpretable symbolic regression

Erwan Le Roux[1], Pierre Tandeo[1,2], Carlos Granero Belinchon[1,2], Melika Baklouti[3], Julien Le Sommer[4], Florence Sevault[5], Samuel Somot[5], Antoine Doury[5], and Mahmoud Al Najar[6]

[1]IMT Atlantique, Lab-STICC, UMR CNRS 6285, Brest, 29238, France
[2]Odyssey, INRIA, IMT Atlantique, IFREMER, CNRS, Rennes, 35042, France
[3]Aix-Marseille Université, Université de Toulon, CNRS, IRD, MIO UM 110, 13288, Marseille, France
[4]Univ. Grenoble Alpes, CNRS, IRD, Grenoble INP, INRAE, IGE, Grenoble, 38058, France
[5]Météo-France, CNRS, Univ. Toulouse, CNRM, Toulouse, 31057, France
[6]INP, IRIT, Université de Toulouse, Toulouse, 31400, France

Correspondence: Erwan Le Roux (erwan.le-roux@imt-atlantique.fr)

Abstract. Projected impacts of climate change are assessed with impact models, such as ecological or hydrological models, driven by climate projections. Uncertainties of projected impacts are estimated by driving impact models with a large ensemble of plausible future climate projections. However, most of the time, this is not possible for practical reasons: computing time, data availability. To fix this issue, we propose an approach that links climate projections to impacts with an interpretable equation. First, this equation is discovered based on simulations of the impact model and their corresponding climate projections. Then, we consider that this equation can be used to emulate the impact model for other climate projections. Specifically, the discovered equation maps each year climate indicators, i.e. a list of yearly and seasonally-averaged climate model variables, to a yearly-averaged impact indicator, i.e. a variable computed from the impact model outputs. In our application, the impact indicator is the annual mean Net Primary Production (NPP) of a risk-relevant regional oceanic area located in the North-Western Mediterranean basin. It is computed from the outputs of a biogeochemical model of the Mediterranean Sea, which is driven by climate projections of a coupled regional climate model of the Mediterranean area. In our methodology, we run nine validation schemes each one providing one equation to predict this impact indicator. Our results show that all discovered equations are linear, even though non-linearity is allowed, and that most of them contain four climate indicators that can be interpreted physically: the sea surface temperatures in winter and spring, the sea surface salinity in spring, and the net downward shortwave flux in winter. Based on these four indicators, we fit a linear equation on the historical period (1986-2005) and the scenario RCP8.5 (2006-2099) that reproduces well the trend and the year-to-year correlation of the impact indicator for the scenario RCP4.5 (2006-2099), which was not used for the fit. The predictions of the linear equation however underestimate the interannual variance of NPP. As a perspective, this equation allows us to approximate the impact indicator at a neglected computational cost, i.e. without running the costly biogeochemical impact model, for any regional climate model outputs available.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

1 Introduction

Future climate change impacts on ecosystems, biodiversity, and human communities are assessed to inform and develop adaptation strategies. These impacts are often evaluated based on the results of modelling chains which comprise three successive steps. The first step involves selecting a socio-economic scenario that describes, among other variables, how greenhouse gas emissions may vary in the future. Then a climate model is run at the global scale for the chosen scenario. Sometimes, this result is downscaled, i.e. refined at a more relevant scale (regional or local), using dynamical or statistical downscaling methods. Finally, the output of the climate model drives an impact model. Such model can assess quantitatively the consequences of climate change on sector-specific systems (e.g. ecosystem, hydrological system) and provide insights on process understanding.

Assessing uncertainties associated with future projections is an issue in climate impact studies (Northrop and Chandler, 2014; Evin et al., 2019). Three main sources of uncertainty are generally accounted for (Hawkins and Sutton, 2009). First, socio-economic uncertainty stems from the uncertain future of global greenhouse gas emission trajectories , and is evaluated with projections from different socio-economic scenarios (Pirani et al., 2024). Second, model uncertainty, also called epistemic uncertainty, is assessed using different climate and impact models, because each model inherently has knowledge gaps. Besides, developers of different models do not make the same assumptions leading to different results from one model to another when forced by the same scenario forcings. Third climate internal variability, which results from the chaotic nature of the climate system, can be estimated using initial-condition large ensembles (Maher et al., 2021).

These uncertainties are usually quantified with a large ensemble of climate projections based on different combinations of scenarios, models, and initial conditions. However, the high computation and storage costs of numerical models can limit the size of the ensemble, especially when high-resolution climate simulations are required to address regional or local climate change impacts. Indeed, climate models require large resources (one regional climate projection taking months on a supercomputer) while costs of impact models range from moderate to expensive (Takakura et al., 2021).

Statistical emulators, also known as surrogate models, can be trained on existing climate projections to mimic the original numerical models in order to produce large ensemble at a much lower computational cost (Doury et al., 2022; Kendon et al., 2025), even though we note that their training costs can vary widely and that they cannot perfectly mimic the original model.

In this study, we rely on statistical emulation to assess socio-economic uncertainty. Specifically, we focus on the task of emulating models based on explored scenarios in order to generate projections for unexplored scenarios. Many approaches have already emulated climate models for this goal. For instance, statistical emulators of global climate models derive temperature and precipitation maps from global emissions using pattern scaling (Santer et al., 1990), simple regression models (Castruccio et al., 2014), time sampling (Tebaldi et al., 2022), causal learning (Boussard et al., 2023), generative models (Bassetti et al., 2024; Bouabid et al., 2025), Gaussian processes, random forests, and custom deep learning architectures (Watson-Parris et al., 2022). For regional climate models, emulators derive high resolution temperature and precipitation maps from large scale atmospheric conditions using principal components (Erlandsen et al., 2020), analogues (Boé et al., 2023), multilayer perceptron or linear regression combined with random forests (Hobeichi et al., 2023), convolutional neural networks (Doury et al., 2022; Bano-Medina et al., 2023; van der Meer et al., 2023), and generative models (Rampal et al., 2024; Addison et al., 2024).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

To go one step further along the climate impact modelling chain, our general objective here is to design statistical emulators of impact models to generate projections for unexplored scenarios. Such emulators should allow for quantitative impact assessments without much prior knowledge (flexibility) and for process understanding (interpretability), as both are often needed for high-stakes decisions. All statistical emulators have a trade-off between flexibility and interpretability. Simpler emulators, e.g. linear regression, are highly interpretable but lack flexibility: they may miss relevant nonlinear relations between input features and target outputs (underfitting). More complex emulators, e.g. neural networks, are flexible but lack interpretability. Such flexible emulators may account too closely for residual variations in the input features, which may lead to overfitting.

Symbolic regression (SR), a.k.a equation discovery, strikes a good balance between interpretability and flexibility. It aims to discover automatically human-readable equations from data (Koza, 1994; Camps-Valls et al., 2023; Makke and Chawla, 2024) by optimizing both their structures and their coefficients, while classical regressions only optimize their coefficients with a fixed structure. In climate science, SR has discovered equations of global temperature (Stanislawska et al., 2012), ocean eddies (Ross et al., 2023), evapotranspiration (Li et al., 2024), cloud cover (Grundner et al., 2024; Beucler et al., 2024b), sea ice albedo (Atmojo et al., 2025) and cost of climate adaptation (McDermott et al., 2025). SR could bridge the gap between data-based approaches and the discovery of parsimonious representations (Balaji et al., 2022; Song et al., 2024; Huntingford et al., 2025).

Our approach consists in discovering with SR an interpretable equation, that generalizes well in unexplored scenarios, in order to explore new scenarios at a lower cost and to assess scenario uncertainties of the climate impact modelling chain. Such an approach has links with the broad challenge of finding climate-invariant equations (Beucler et al., 2024a). In our application, we discover an equation using a 115-year long simulation for training, covering a past historical period and a future scenario under high emission (RCP8.5), and we evaluate the predictive capacity of this equation with an intermediate scenario (RCP4.5) simulation unseen during the training (Fig. 1).

This study focuses on marine biogeochemistry, where statistical emulators of biogeochemical models are seldomly used (Mattern et al., 2013; Hemmings et al., 2015; Schartau et al., 2017; Skákala et al., 2023; Buchanan et al., 2025). Here, we emulate an impact indicator defined as the annual mean Net Primary Production (NPP), in grams of carbon per year, summed over an offshore area off the Gulf of Lion located in the North-Western Mediterranean basin. NPP is the total rate of organic carbon production by photosynthesis of marine phytoplankton minus their respiration (Sigman and Hain, 2012).

Phytoplankton plays an invaluable role: it is at the basis of marine food webs, controlling the energy available to upper trophic levels, and it produces around 50% of the oxygen on Earth. Climate change is expected to impact phytoplankton. Indeed, the expected decrease of vertical mixing in the ocean, which brings nutrients to the photic zone during winter (surface layer of the ocean that receives sunlight), may reduce primary production (Mora et al., 2013; Hutchins and Tagliabue, 2024; Pagès et al., 2020). Between 1990s and 2090s, a global reduction of NPP is expected ranging from -2.0% to -8.6% under RCP2.6 (low emissions scenario) and RCP8.5 (very high emissions scenario), respectively (Bopp et al., 2013). In the western Mediterranean basin, projections of primary production remain uncertain: some studies expect a decrease (Macias et al., 2015; Richon et al., 2019; Pagès et al., 2020) while others expect an increase (Herrmann et al., 2014; Moullec et al., 2019). Therefore, a better characterization of the various sources of the projection uncertainty seems required to hope to deliver more robust

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

results concerning the future evolution of phytoplankton under a changing climate at this regional scale. Our study targets to contribute to this general long-term scientific objective with application outside the biogeochemical modelling field.

This paper is organized as follows. Section 2 presents our data. Section 3 explains our methodology. Results, discussion and conclusions are introduced in Sects. 4, 5 and 6 respectively.

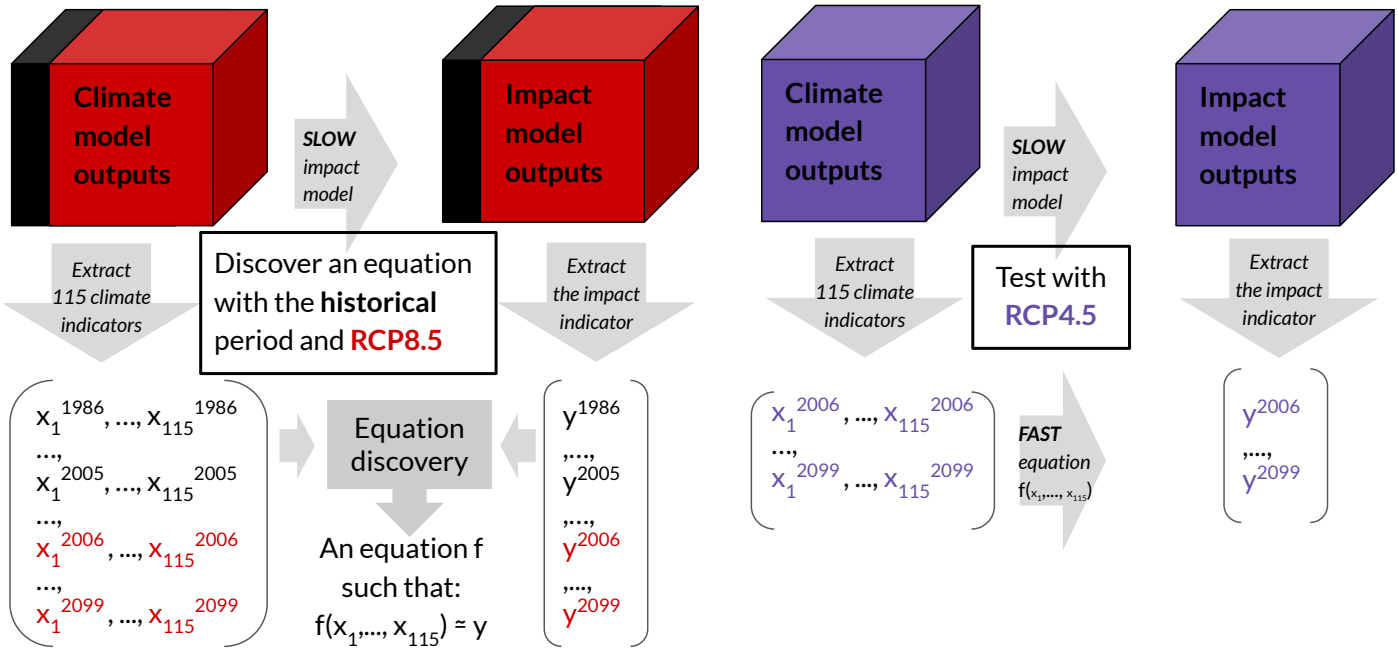

Figure 1. Overview of our approach. Left: an equation, mapping each year climate impact indicators to the impact indicator, is discovered with the historical period and the scenario RCP8.5. Right: the equation predicts the impact indicator for the scenario RCP4.5.

2 Data

In this study, we focus on an offshore area off the Gulf of Lion (Fig. 2) in the North-Western Mediterranean basin. This gulf is one of the most productive areas in the Mediterranean Sea (Bosc et al., 2004) in terms of primary production: the production of organic matter by phytoplankton photosynthesis (Sigman and Hain, 2012). This productivity is fueled by nutrients coming from winter vertical mixing and high discharge of the Rhone river (Auger et al., 2011, 2014; Many et al., 2021). We use a coupled regional climate model (Sect. 2.1) on a historical period (1986-2005) and two RCP scenarios (2006-2099) to drive a biogeochemical model of the Mediterranean Sea (Sect. 2.2). Based on the outputs of both models, climatic impact-drivers and indicators (Sect. 2.3) are computed in the studied region.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

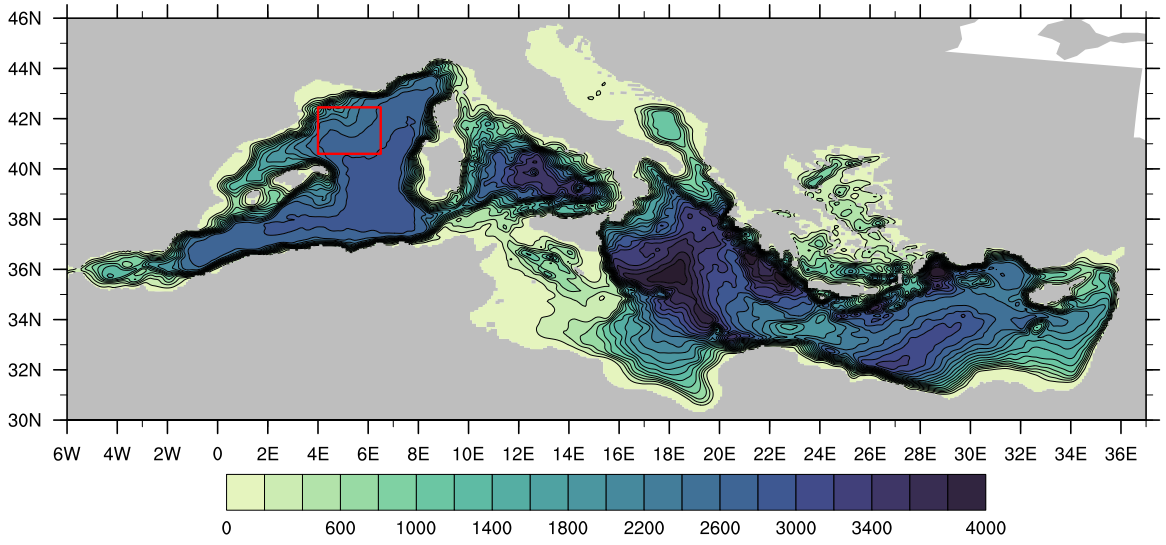

Figure 2. Ocean bathymetry (in m) of the regional climate model CNRM-RCSM4 over the Mediterranean Sea. In this study, we focus on an offshore area off the Gulf of Lion (highlighted in red) which is located in the North-Western Mediterranean basin.

2.1 The regional climate model CNRM-RCSM4

CNRM-RCSM4 is a fully-coupled regional climate model dedicated to the study of the Mediterranean climate and sea (Sevault et al., 2014). It simulates at high-resolution the main components (atmosphere, ocean, land, rivers) of the Mediterranean regional climate system and their interactions at high-frequency (daily). The ocean component of the model that will be mostly used in the study has a horizontal resolution of $1/8^o$ (6 to 12 km), a vertical resolution of 43 vertical levels (levels ranging from 6 m at the surface to 200 m, see Fig. 2), and a time step of 20 minutes. Lateral boundary conditions for the atmosphere and the Atlantic ocean comes from the general circulation model CNRM-CM5 (Voldoire et al., 2013), one of the GCM involved in CMIP5 (5th generation of the Coupled Model Intercomparison Project). Ocean boundary conditions were de-trended from the 152-year control run trend, and un-biased using the 1960-2005 temperature and salinity means of the NEMOVAR-COMBINE reanalysis (Balmaseda et al., 2010). Simulations are run for two periods: a historical period (1986-2005) during which the forcings (e.g. greenhouse gas concentrations) are based on observations and a projection period (2006-2099) in which the forcings follow socio-economic prospective scenarios. Projections are available for two scenarios: a moderate (RCP4.5) and a high (RCP8.5) emission scenario. For details on this model, we refer to Sevault et al. (2014). CNRM-RCSM4 is one of the coupled regional climate model contributing to the international initiative Med-CORDEX (Ruti et al., 2016), Evaluation of the historical simulation and description of the future projections can be found in Darmaraki et al. (2019); Soto-Navarro et al. (2020); Parras-Berrocal et al. (2024). In addition, the historical and RCP8.5 simulations have been already used to drive impact models (Moullec et al., 2019; Pagès et al., 2020; Fabri-Ruiz et al., 2024).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

2.2 The biogeochemical model Eco3M-MED

Eco3M-MED is an ocean biogeochemical model mainly dedicated to the Mediterranean Sea (see Baklouti et al., 2021, and references within). This model is based on differential equations describing mechanistically fluxes and transformations of biogeochemical components. Several organic components are accounted for: two phytoplankton compartments (large, small), three zooplankton compartments (copepods, ciliates, heterotrophic nanoflagellates), a compartment of heterotrophic bacteria, and a pool of dissolved organic matter. Inorganic components are also represented: nutrients (nitrate, ammonium, phosphate), as well as two compartments of detrital particulate matter (small and large particles). Organisms are represented in terms of biomass and abundance, which makes it possible to regulate process rates through intracellular quotas and ratios of biogenic elements. The Eco3M-MED has been driven by the regional climate model CNRM-RCSM4 for the historical period (1986-2005) and the two RCP projections (RCP4.5 and RCP8.5 over 2006-2099). For more details about the evaluation on the historical period, we refer to Pagès et al. (2020).

2.3 Climate indicators and the impact indicator

Here, an indicator denotes the spatio-temporal average of a variable. Spatially, all indicators are averaged over the whole study area and the entire water column. Temporally, the impact indicator, the net primary production (NPP), is averaged over a year starting in December, because it marks the start of a new "biogeochemical year" as winter mixing often induce a "reset of the biogeochemical memory" in the surface layer. For instance, the NPP in 2000 is the average from the 1st of December 1999 to the 31st of November 2000. Climate indicators are averaged yearly and seasonally (winter, spring, summer, autumn). In total we have 115 climate indicators: each one of all different variables of CNRM-RCSM4 (23 variables in total) are averaged yearly and seasonally. In Figure 3, we illustrate one climate indicator (sea surface temperature in spring) and the impact indicator of our application (annual net primary production) for the historical period (1986-2005) and the two RCP scenarios (2006-2099).

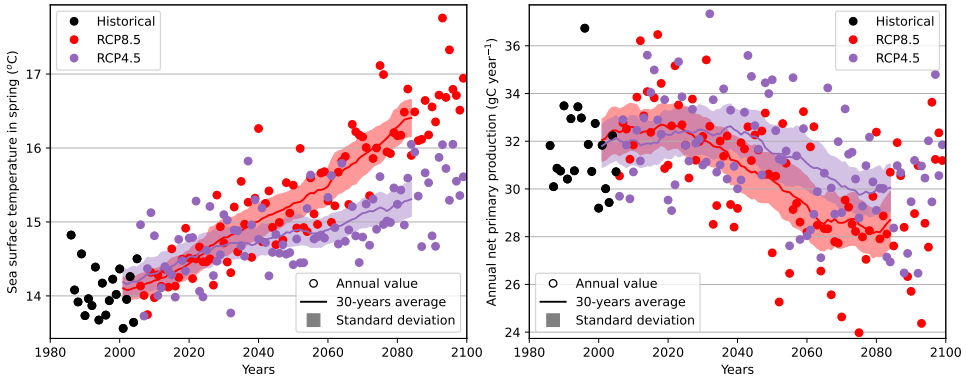

Figure 3. Left: an example of a climate indicator (sea surface temperature in spring in the study area). Right: the impact indicator of our application (annual net primary production in the study area). Colored dots correspond to their values for the historical period (1986-2005) or two RCP scenarios (2006-2099). Colored lines show averages with a sliding window of 30 years, while standard deviation is shaded.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

3 Methodology

3.1 Symbolic regression

SR is a regression in a space $\mathcal{F}$ of equations with a simple input–output relationship (Camps-Valls et al., 2023). This equation space $\mathcal{F}$ is often defined by the combination of primitive operations. For instance, if we combine together the operations $+$, $-$, and $\times$ then $\mathcal{F}$ contains polynomials of arbitrary degree in x such as $2 \times x^2 - x + 1$. In other words, SR optimizes the structure of the equation, which is composed of primitive operations. Nevertheless, due to the large space $\mathcal{F}$, SR is an NP-hard problem (Virgolin and Pissis, 2022), i.e. some instances of the SR problem cannot be solved in polynomial time. This is the reason why existing SR algorithms are generally based on heuristics, such as genetic algorithms, that do not guarantee optimality. In the literature, there exists two main classes of approaches to find approximate solutions. Discrete search methods rely on tree-based representation of equation to search for optimal equations using heuristics such as genetic programming (Koza, 1994; Schmidt and Lipson, 2009), mixed-integer nonlinear programming (Cozad and Nikolaos, 2018; Cornelio et al., 2023), physics inspired strategies (Wu and Tegmark, 2018; Udrescu and Tegmark, 2019), reinforcement learning (Petersen et al., 2019) or Markov Chain Monte Carlo (Guimerà et al., 2020). Continuous search methods use sparse regression to solve a relaxation of SR based on a large but fixed library of equations defined manually (McConaghy, 2011; Brunton et al., 2016) or within a neural network (Martius and Lampert, 2016; Sahoo et al., 2018). In this study, we rely on a discrete search method called PySR (Cranmer, 2023), due to its flexibility and high success rate in benchmarking tests (de Franca et al., 2024). For more details on PySR, we refer to App. A.

3.2 Interpretable symbolic regression

Most SR methods rely on the length of the equation as a proxy for interpretability, even if the assumption that shorter equations are more interpretable is not always valid (Yu et al., 2025). Indeed, short equations with complex internal structure, i.e. containing nested equations and/or terms with multiple variables, can be difficult to analyze and interpret.

In this study, we only consider interpretable equations, i.e. without nested functions and/or cross terms, by restricting with the PySR library the space of allowed equations $\mathcal{F}$. Let x_1, ... x_n denote the features, i.e. potential variables of the equation $f(x_1, ...x_n)$. In Equation 1, we define the restricted space of equations $\mathcal{F}$ as the linear combination of interpretable terms from the set $\mathcal{T}$, i.e. terms with at most one variable x_i and where the function inverse, square and square root are at most used once:

$$\mathcal{F} = \{\lambda_1 \times t_1 \ + \ ... \ + \ \lambda_l \times t_l \,|\, l \in \mathbb{N}, \lambda_1 ... \lambda_l \in \mathbb{R}, t_1 ... t_l \in \mathcal{T}\} \quad \text{where} \quad \mathcal{T} = \bigcup_{i=1}^{n} \{x_i^p | p \in \{-2, -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, 2\}\} \tag{1}$$

For example, $2x_1 + 3\sqrt{x_2} \in \mathcal{F}$ and $4x_1^2 + \frac{5}{x_2} \in \mathcal{F}$ and $\frac{6}{\sqrt{x_1}} + \frac{7}{x_2^2} \in \mathcal{F}$, while $x_1 \times x_2 \notin \mathcal{F}$ and $x_1^3 \notin \mathcal{F}$ and $x_1^{1/4} \notin \mathcal{F}$.

Any equation from this restricted space $\mathcal{F}$ can be easily analyzed. For instance, we can assess that the variable x_i in the term $\lambda \times x_i^p$ has a positive contribution if $\lambda > 0$, while it has a negative contribution if $\lambda < 0$.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

3.3 Validation schemes

Each validation scheme requires two sets (train set, validation set) and consists of two steps. First, 500 sets of hyperparameters of PySR are sampled randomly (App. B) and for each set of hyperparameters, one equation $f \in \mathcal{F}$ is discovered on the train set using interpretable symbolic regression with the PySR library (Sect. 3.2). Then, among these 500 equations, we select only the equation f^* minimizing the prediction error on the validation set, i.e. the root mean squared error between the predicted impact indicator and the reference impact indicator. In this article, f^* is referred to as the validated equation, and its performance is evaluated on our test set: the scenario RCP4.5.

In total, we run 9 different validation schemes. For every scheme, the discovery period (historical period and RCP8.5) is split between a train set and validation set. The 9 different schemes result from the combination of 3 different splitting strategies and 3 sizes for the validation set. On the one hand, the validation set represents either 20%, 25% or 30% of the discovery period. On the other hand, the validation set is either chosen at the center of the discovery period (around 2042), sampled randomly, or sampled randomly based on quantiles. For the latter method, datapoints of the discovery period are first sorted based on the values of the impact indicator, then split into bins, and finally one validation datapoint is sampled from each bin.

4 Results

4.1 Overview of validated equations

Table 1 shows the validated equations, i.e. the equations selected for each of the 9 validation schemes. These 9 equations contain only linear terms, even if other operations are allowed (Eq. 1) such as square, square root, inverse. Given enough samples for the random search, we would expect these 9 equations to converge toward a unique mechanistic equation. However, we observe that all discovered equations are different.

Figure 4 illustrates the prediction performance (on the validation and test sets) for the nine validated equations of Table 1. It also shows the performance of the equation discovered using interpretable symbolic regression with the default set of hyperparameters of PySR. For each validation scheme, as expected, we observe that the prediction error on the validation set is always smaller for the validated equation (optimized on this set) than for the baseline equation. On the test set, we find that six times out of nine, the prediction error is smaller for the validated equation. The best prediction error on the test set (root mean squared error equal to 1.59) is obtained with the validated equation of the validated scheme "20% selected by quantile binning". Overall on the test set, the performance of the validated equation is often not significantly better than the baseline (three validated equations perform worse than a baseline). This is likely due to the fact that validation sets, which contain between 23 (20%), and 35 (30%) datapoints, are always too small. In the literature, validation schemes are rarely used to discover equations. Here, we show that such scheme can in a majority of cases lead to improved performance than baseline, but is not guaranteed to. For the 3 validation sets "at the middle", i.e. when the equation is optimized to interpolate (predict years with a moderate warming based on years with a low and high warming), the performance on the test is always better than the baseline.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Validation set	Validated equation, i.e. Equation selected by the validation scheme
20% of random	$-34\text{MWS}_{\text{Annual}} + 9\text{SSS}_{\text{Annual}} - 3\text{SST}_{\text{MAM}} + 15\text{SSSH}_{\text{MAM}} + 001\text{SP}_{\text{MAM}} - 03\text{TH}_{\text{JJA}} + 261097\text{upVelo}_{\text{MAM}} + 452$
20% of quantile with binning	$-36\text{MWS}_{\text{MAM}} + 6\text{SSS}_{\text{MAM}} - 1\text{SST}_{\text{DJF}}$
20% of middle	$12\text{SSS}_{\text{MAM}} - 1.8\text{SST}_{\text{DJF}} - 2\text{SST}_{\text{MAM}} + 0.1\text{Shortwave}_{\text{DJF}} + 5\text{StratIndex}_{\text{MAM}} + 789$
25% of random	$-38\text{MWS}_{\text{Annual}} - 20\text{MWS}_{\text{MAM}} + 6\text{SSS}_{\text{Annual}} - 1\text{SST}_{\text{Annual}} + 157\text{ZonCur}_{\text{SON}}$
25% of quantile with binning	$8\text{SSS}_{\text{MAM}} - 2\text{SST}_{\text{DJF}} - 1.8\text{SST}_{\text{MAM}} + 21\text{SSSH}_{\text{DJF}} + 0.1\text{Shortwave}_{\text{DJF}} + 888$
25% of middle	$10\text{SSS}_{\text{MAM}} - 3\text{SST}_{\text{DJF}} - 2\text{SST}_{\text{MAM}} + 0.1\text{Shortwave}_{\text{DJF}} + 5\text{StratIndex}_{\text{DJF}} + 925$
30% of random	$7\text{SSS}_{\text{MAM}} - 1\text{SST}_{\text{MAM}}$
30% of quantile with binning	$8\text{SSS}_{\text{MAM}} - 1\text{SST}_{\text{MAM}} + 0.1\text{Shortwave}_{\text{DJF}}$
30% of middle	$-4\text{NSHF}_{\text{MAM}} + 9\text{SSS}_{\text{MAM}} - 4\text{SST}_{\text{DJF}} + 0.1\text{Shortwave}_{\text{DJF}} + 4\text{StratIndex}_{\text{DJF}} + 880$

Table 1. Equations selected for the 9 different validation schemes. Each variable in the equation is written as $\text{VariableName}_{\text{Period}}$ where VariableName represents a short name for the climate variable (App. C) and Period is the season when the average is computed, which can be Annual, winter (DJF), spring (MAM), summer (JJA), or autumn (SON). SSS and SST respectively refer to sea surface salinity and temperature. Shortwave refers to the net downward shortwave flux at the sea surface, MWS to the meridional wind stress, Stratindex to the stratification index, TH to the net downward total heat flux, SSSH to the sea surface steric height, SP to the surface pressure, upVelo to the upward velocity, ZonCur to the zonal current, NSHF to the net surface heat flux. For more details on the physical unit we refer to App. C.

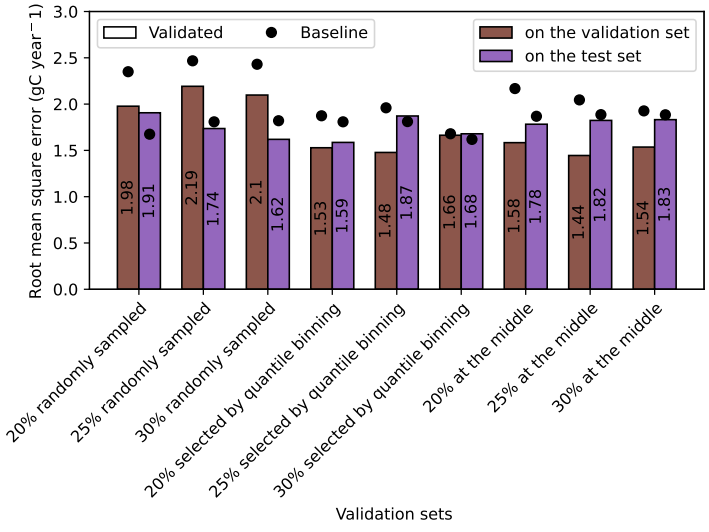

Figure 4. Prediction performance for the 9 validation schemes. For each validated scheme, the root mean squared error of the validated equation on the validated set (between 20% and 30% of the historical period and scenario RCP8.5) and test set (scenario RCP4.5) are displayed with brown and purple bars, respectively. For the baseline equation, discovered using interpretable symbolic regression with the default set of hyperparameters of PySR, the root mean squared error is displayed with a black dot.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

4.2 Analysis of validated equations

The proposed definition of interpretable equation (Sect. 3.2) makes it possible to directly analyze any discovered equation. Due to the lack of convergence to a unique mechanistic equation, we decided to analyze together all the equations found for the 9 different validation schemes (Tab. 1). In practice, we observe that these equations are linear even though they could have contained more complex terms. Thus, any equation f can be rewritten as $f(\mathbf{x}) = f(x_1, ..., x_{135}) = \lambda_0 + \sum_{i=1}^{135} \lambda_i x_i$, where $\lambda_i = 0$ if the climate indicator x_i does not appear in the equation. If $\lambda_i > 0$ the climate indicator x_i has a positive contribution to the predicted value, while it has a negative contribution if $\lambda_i < 0$. Furthermore, for any equation f and input $\mathbf{x}_y$ for the year y, we can compute the relative contribution of the i-th climate indicator as $RC_i(f, \mathbf{x}_y) = 100 \times \left| \frac{\lambda_i x_i}{f(\mathbf{x}_y)} \right|$, which indicates the percentage of absolute contribution to the predicted value. In practice, we consider the relative contribution $RC_i^{1986-2099}(f)$ for the historical period and scenario RCP8.5 (1986-2099), i.e. $RC_i^{1986-2099}(f) = \frac{1}{114} \sum_{\mathrm{y}=1986}^{2099} RC_i(f, \mathbf{x}_y)$.

In Figure 5, bars illustrate the contribution of climate indicators based on the 9 equations of Table 1. For each indicator x_i, the color of the bar shows $RC_i^{1986-2099}(f)$ averaged on all equations f where the indicator is involved. Then, the rate of equations where the indicator is involved (out of 9) is represented with the length of the vertical bar. Finally, the direction of the bar (up or down) highlights the sign (positive or negative) for the contribution of the indicator to the predicted value.

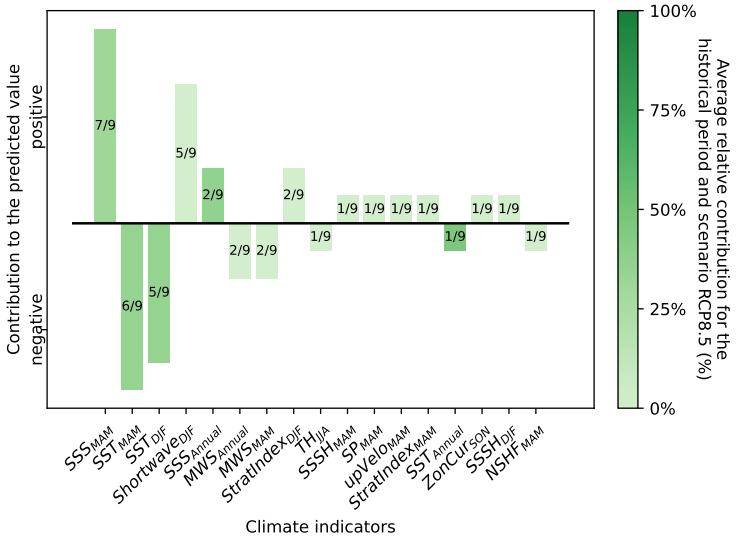

Figure 5. Contribution of climate indicators based on the 9 equations of Table 1. Each climate indicator is associated with a vertical bar. Its length corresponds to the rate of equations where this indicator is involved, while its direction (below or above a horizontal black line) indicates whether the contribution of this indicator is positive or negative. The color of the bar shows the relative contribution for the historical period and scenario RCP8.5, i.e. $RC_i^{1986-2099}(f)$, averaged on all equations f where the indicator is involved. SSS and SST respectively refer to sea surface salinity and temperature. Shortwave refers to the net downward shortwave flux at the sea surface, MWS to the meridional wind stress, Stratindex to the stratification index, TH to the net downward total heat flux, SSSH to the sea surface steric height, SP to the surface pressure, upVelo to the upward velocity, ZonCur to the zonal current, NSHF to the net surface heat flux.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

We find that four climate indicators are in a majority of the 9 equations of Table 1. On the one hand, SST_{MAM} and SST_{DJF}, the sea surface temperature in spring and in winter, have both a negative contribution to the predicted value and an average relative contribution roughly equal to 35%. On the other hand, SSS_{MAM} the sea surface salinity in spring and $Shortwave_{DJF}$, the net downward shortwave flux in winter, have both a positive contribution but with a different average relative contribution: 30% for SSS_{MAM} and less than 1% for $Shortwave_{DJF}$. For a visualization of these indicators, we refer to App. E.

4.3 Final equation based on validated equations

Our objective is to find a single equation mapping climate indicators to the impact indicator and generalizing well in unexplored climate scenarios. However, different validation schemes lead to different validated equations, and do not lead to the discovery of a unique mechanistic equation (Sect. 4.1). After analyzing validated equations (Sect. 4.2), we observe that linear equations are always preferred and that four climate indicators are considered in most validated equations: SST_{MAM}, SST_{DJF}, SSS_{MAM} and $Shortwave_{DJF}$. Therefore, we choose to rely on a linear equation with these four identified climate indicators. We fit this equation on the historical period (1986-2005) and the scenario RCP8.5 (2006-2099) and obtain:

$$\text{Annual net primary production} = 6.2\text{SSS}_{\text{MAM}} - 0.086\text{SST}_{\text{DJF}} - 1.03\text{SST}_{\text{MAM}} + 0.11\text{Shortwave}_{\text{DJF}} + 106 \quad (2)$$

Qualitatively, this equation can be interpreted physically. SST_{MAM} and SST_{DJF} have both a negative contribution to the annual net primary production. Indeed, high sea surface temperatures are generally accompanied by more pronounced stratification of the surface layer, which reduces the efficiency of vertical mixing in the ocean. Yet this vertical mixing is the main source of nutrient input into the photic zone far from the coast, and therefore has a significant influence on phytoplankton growth and primary production. This explains, at least in part, the negative correlation between primary production and SST. Furthermore, vertical mixing cools the surface waters, and very cold surface waters generate convection and mixing. Both conditions favor primary production, and provide another explanation for the negative correlation between SST and primary production. SSS_{MAM} has a positive contribution because saltwater is denser and promotes convection and vertical mixing, which facilitates the transport of nutrients from the depths to the surface and supports primary production. Finally, $Shortwave_{DJF}$ has a positive contribution, as solar energy is generally the factor that most limits photosynthesis and primary production in winter.

Quantitatively, we provide various metrics to diagnose the quality of this equation. In Figure 6, we compare reference values of our impact indicator against values predicted by the equation for both the discovery period and the test period. All metrics (root mean squared error, mean relative average error, detrended correlation) are comparable between the discovery period (historical period and scenario RCP8.5) and test period (scenario RCP4.5). Root mean squared error and mean relative average error are slightly smaller for the scenario RCP4.5. This is likely because this scenario contains less extreme values, which can be hard to predict. For instance, six reference values are below 26 gC year^{-1} for RCP8.5 (Fig. 6 **a**), while no values are below this threshold for RCP4.5 (Fig. 6 **b**). Overall, for both scenarios, every absolute error remains below 5 gC year^{-1}, or below 20% in terms of absolute relative error. Climatologically, we can compute the trend in average annual net primary production between an historical period (1986-2005) and a future period (2080-2099) using a difference between the average value of

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

these two periods. We observe that the decreasing trend is overestimated for RCP8.5 (predictions gives -11.78% while the reference is -7.2%) and underestimate for RCP4.5 (predictions gives -4.96% while the reference is -6.13%). In Figure 6 **c** and Figure 6 **d**, we illustrate the 30-years average computed with a moving time window of 30-years. We observe that the 30-years average of the prediction remains close to the 30-years average of the reference. Overall, our results show that the chosen equation reproduces well the trend of the impact indicator over the period 2006-2100 for the scenario RCP4.5. The chosen equation also reproduces well the year-to-year correlation (YC) for the scenario RCP4.5: YC equals 0.35 for the reference and 0.4 for the prediction. However, YC is largely overestimated for the historical period and scenario RCP8.5: YC is 0.48 for the reference and 0.74 for the prediction.

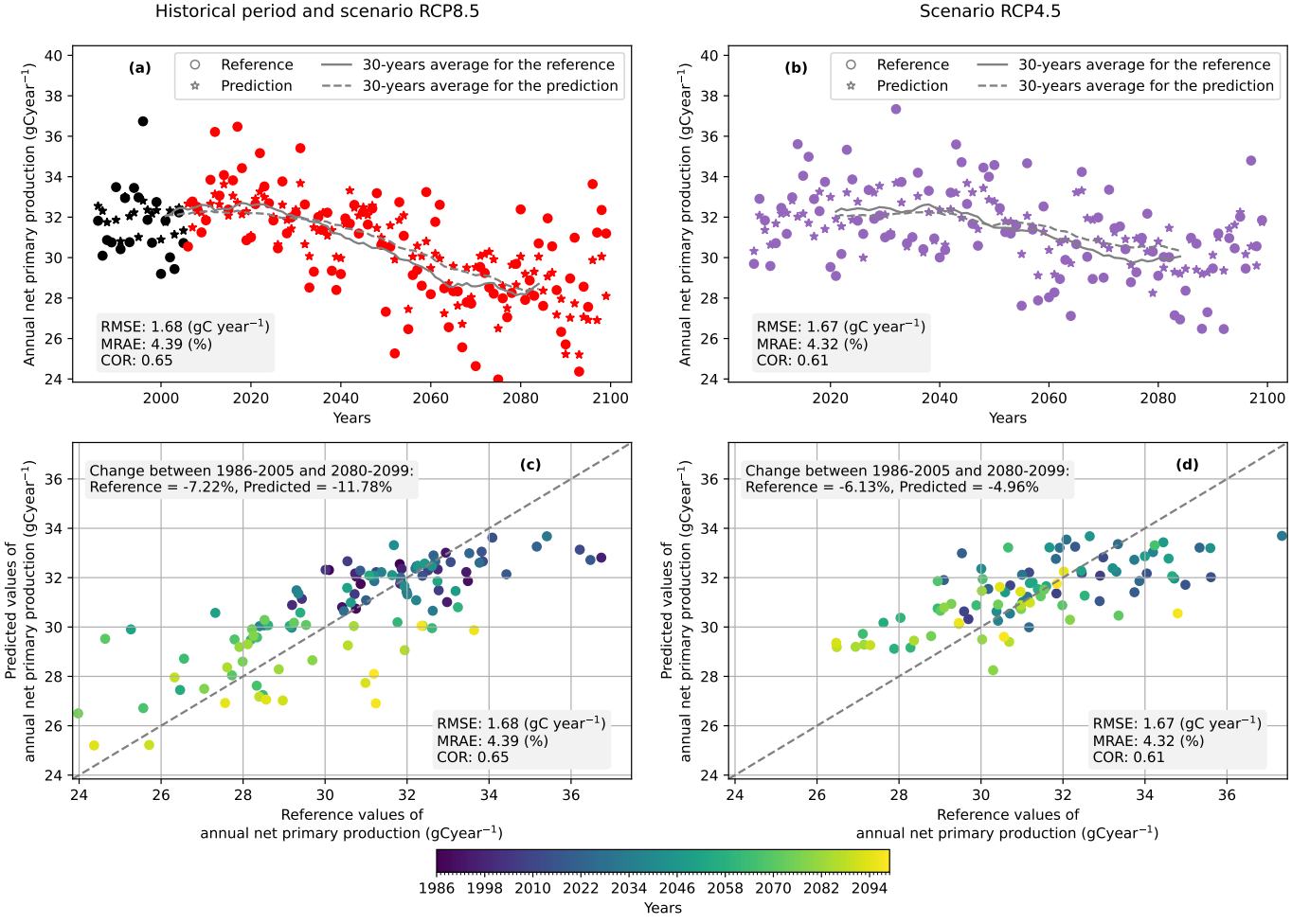

Figure 6. The comparison of predicted values by Eq. 2 with the reference values is illustrated with **(a, b)** time series plots and **(c, d)** scatter plots. This comparison is provided for two periods: **(a, c)** the historical period and scenario RCP8.5, and **(b, d)** the scenario RCP4.5. Several metrics are displayed: the root mean squared error (RMSE), the mean relative average error (MRAE), the detrended correlation (COR).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

5 Discussion

5.1 Data

In this study, numerical models (Sect. 2.1 and 2.2) are run at the scale of the whole Mediterranean Sea, while climate indicators and the impact indicators are aggregated over an offshore area off the Gulf of Lion. Practically, we choose to focus on such area because it seems too difficult to find an equation, based on climate indicators at the scale of the whole Mediterranean Sea, that can predict the aggregated primary production of regions with widely different biogeochemical context (D'ortenzio and Ribera Dalcaì, 2009). Thus, we focus on one of the most productive areas in the Mediterranean Sea where the biogeochemical mechanisms are likely more homogeneous, even if studying only such a restricting area could have its limits as changes in primary production can be driven by internal production as well as by external production advected inside the area. Thus, as a potential extension, primary production for the whole Mediterranean Sea could be estimated by discovering one equation for each of its different bioregions. In detail, here we did not focus on the Gulf of Lion, but on an offshore area (Fig. 2) that excludes the coast where physical and biogeochemical phenomena are more complex. To study such coastal areas, we would need both a more detailed spatial resolution and a precise account for nutrients coming as runoff (set as a monthly climatology in the biogeochemical simulations), as well as consideration of interactions with the ocean floor.

5.2 Methodology

In this article, we propose to discover an "input-output" equation (Song et al., 2024), mapping each year climate indicators to the impact indicator. To our knowledge, this approach is a promising innovation that could pave the way to develop simple emulators for other climate impact studies. However, here, we implicitly assume that impact indicators of two successive years are independent. For the primary production, it is not a far-fetched hypothesis, because winter mixing may induce a "reset of the biogeochemical memory" in the surface layer. In particular, the studied area is known to have the most intense mixing in the Mediterranean Sea. This is the reason why we extract indicators on years that start in December at the beginning of the winter (Sect 2.3). In practice, we note that other starting months could have been considered, e.g. September, without changing the rest of our methodology. It is worth noting that the winter mixing is expected to decrease in the Gulf of Lion with global warming (Somot et al., 2006; Parras-Berrocal et al., 2024). In practice, this independence assumption is likely exaggerated because the Pearson correlation equals 0.48 between impact indicators of two successive years, which indicates a moderate correlation. For other impact variables, this hypothesis (independence between two successive impact indicators) may not hold, which limits the range of applicability of our approach. If successive impact indicators are highly dependent, one potential solution would be to learn an equation representing the dynamic. For instance, Berthold et al. (2025) rely on another symbolic regression (SR) library called SINDy (Brunton et al., 2016) to identify environmental drivers involved in differential equations that represent how phytoplankton bloom patterns evolve. Interestingly, in the same way than our approach in Sect. 4.2, this study also relies on SR to quantify the contribution (positive or negative) of variables to the target (see Fig. 3 and Tab. 1 of Berthold et al. 2025).

Otherwise, SR methods are often designed for large amounts of data, i.e. at least a few hundred to a few thousand data points (Camps-Valls et al., 2023). Here Eq. 2 is fitted on 114 datapoints: the historical period (1986-2005) and the projection period

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

(2006-2099). For future application, to have more datapoints, we advise to start the historical run in 1950. Preliminary results, with other machine learning models, suggested that the most extreme scenario (RCP8.5 in our case) should be preferred for discovering the equation, likely because it spans a wider range of climate states.

Our final equation (Eq. 2) is a linear equation with four variables. Such an equation could have probably been obtained with simpler approaches, e.g. linear regression combined with feature selection or with L1 regularization. However, the advantages of our approach is that we do not assume any linearity, and allow also for non-linear terms in the final equation.

5.3 Results

SR is viewed as a tool that discovers physical law from a few hundred to a few thousand data points with 1-10 input variables (Camps-Valls et al., 2023). In the case of smaller datasets with hundreds of variables, we suggest that it should be viewed as an empirical tool aiming at approximating an equation as good as possible. Indeed, in Table 1 we observe that all equations are different. This diversity of discovered equations, i.e. lack of convergence, is because no equation can predict exactly the impact indicator based on chosen climate indicators, which is likely due to the small number of datapoints for the validation schemes. The relatively small number of iterations (between 10 and 1000 see App. A) could also explain this lack of convergence. In practice, in our early experiments, large numbers of iterations often led to overfitting due to the small number of datapoints. This is the reason why we restricted the number of iterations around the default number of iterations for PySR (100 iterations), and why the small number of datapoints may be the root cause for the diversity of discovered equations.

In our results, we exemplify how SR can be used to identify four climate indicators that correlates well with the impact indicator. Two of these four indicators are in winter, while the other two are in spring. This result matches well the seasonal repartition of annual net primary production in the studied area which indicates that more than 60% of the production occurs in winter and spring (App. D). More generally, this result might indicate that, in the studied area, annual primary production does not depend only on winter preconditioning (nutrient supply through winter mixing), but also on spring conditions (production seems to correlate well with the salinity and temperature at the sea surface). Thus, equation discovery makes it possible to both validate existing knowledge (winter preconditioning) and provide novel links (spring conditions also influence the bloom).

Our final equation (Eq. 2) is directly interpretable physically. SST_{MAM} and SST_{DJF} are negatively correlated with the annual net primary production, while SSS_{MAM} and $Shortwave_{DJF}$ are positively correlated. As explained in Sect. 4.3, the explanation for $Shortwave_{DJF}$ is straightforward because it is the factor that most limits photosynthesis and primary production in winter. Besides, SST_{MAM} and SST_{DJF} are likely negatively correlated (positively correlated for SSS_{MAM}) with an intense vertical mixing of the ocean, which brings nutrients to the photic zone. However, correlation does not imply causation. Thus we cannot conclude whether each of these factors are a precursor or a result of an intense vertical mixing. Overall, we were surprised that the mixed layer depth (MLD), i.e. a variable that characterizes the vertical mixing, was not selected in any validated equations (Tab. 1). Thus, we could assume that here the SST and SSS indirectly integrate the MLD.

In Figure 6, we show that the predicted 30-years average matches well with the reference 30-years average. This is likely because the equation is found by optimizing the root mean squared error (the default loss for the library PySR). The main drawback of such optimization is that extreme values are often not well predicted by the equation. Specifically, in our case,

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

we observe that the linear equation underestimates the interannual variance of the reference dataset (Fig. 6). Thus, for studies aiming at predicting well extremes or the interannual variance, we would advise to change the optimization loss of PySR.

6 Conclusions

We propose an application of automatic equation discovery, also known as symbolic regression (SR), for climate impact. Specifically, we describe a novel task: learning with a single scenario an equation that maps each year climate indicators (list of yearly and seasonally-averaged climate model variables) to an impact indicator (a scalar computed from the outputs of an impact model). Our contribution is to propose methodological guidelines, and to illustrate them with a case study.

For the offshore area considered in our study, we show that equation discovery makes it possible to validate existing knowledge, such that the importance of winter mixing as a precondition for spring bloom, and to understand novel links: spring conditions also influence the bloom, sea surface temperature and salinity can be used to approximate the annual production.

Many potential extensions of this work could be envisioned. First, our methodology should integrate a quantitative assessment of its predictive uncertainty, as such information is needed for high-stakes decisions. Then, we could extend our approach to emulate a chain of several models such as climate impact chains which are composed of three models: a global climate model, a regional climate model, and an impact model. Finally, our methodology assumes that impact indicators of two successive years are independent. This assumption limits its range of applicability. Instead, a potential solution would be to learn the dynamic, i.e. discover the ordinary differential equation that drives the impact indicators. Therefore, the challenge would be to adapt existing equation discovery approaches for dynamical systems (Bongard and Lipson, 2007; Schmidt and Lipson, 2009; Brunton et al., 2016) to our data-scarce context.

Author contributions. ELR, PT, CGB, MB, JLS, SS designed the research. ELR performed the analysis and drafted the first version of the manuscript. MB and FS supplied the data. All authors discussed the results and edited the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Data availability. Datasets and source code used in this article are available at https://github.com/lerouxerwan/EquationDiscoveryForClimateImpact

Acknowledgements. This study has received funding from Agence Nationale de la Recherche under Grant ANR-22-POCE-0003 (PPR-OCEAN MEDIATION project) funded by the French Government "France 2030" program of the French National Research Agency (ANR). The CNRM-RCSM4 simulations used in this study are part of the Med-CORDEX international initiative (https://med-cordex.github.io/).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Appendix A: More details on the library PySR for symbolic regression

Let $D = \{(\mathbf{x}^{(i)}, y^{(i)}) | i = 1, ..., \mathrm{m}\}$ represent a regression dataset with m samples, where $\mathbf{x}^{(i)} \in \mathbb{R}^{\mathrm{n}}$ are the features for the i-th sample, and $y^{(i)} \in \mathbb{R}$ is the target. The objective of SR is to determine an equation $f^* \in \mathcal{F}$ that strives for an optimal balance between accuracy, characterized by the empirical error $l(f) = \frac{1}{m}\sum_{i=1}^{m}(f(\mathbf{x}^{(i)}) - y^{(i)})^2$, and complexity $c(f)$. In statistical learning, $c(f)$ generally corresponds to L2 or L1 regularization, i.e. estimated coefficients of the equation f are constrained towards zero in order to promote sparsity: f should be as simple as possible but not simpler. By contrast, for the task of SR, the complexity $c(f)$ can have many different definitions (Smits and Kotanchek, 2005; Kommenda et al., 2015; Udrescu et al., 2020). The library PySR has an intuitive definition of complexity: each coefficient, variable and operator in the equation f counts as 1 of complexity. In other words, if we represent the equation as a tree, then complexity equals the number of nodes in the tree.

PySR is based on an algorithm that iteratively builds a Pareto-optimal set of equations containing the most accurate equation at each complexity. Specifically, it proceeds in two steps: i) estimation of a Pareto-optimal set of equations $\{f_{C_1}, f_{C_2}, ..., f_{C_P}\}$ where each equation f_{C_i} minimizes the empirical error for the complexity C_i ii) selection of a single equation from this set. In Figure A1, we illustrate a Pareto-optimal set of equations $\{f_1, f_3, f_5, f_7, f_9\}$ estimated with PySR on data generated with the equation $x^2 + 2 \times x + 3$. We observe that the empirical error decreases with complexity, until the complexity 9 where the error reaches zero because PySR finds the equation that generated the data. Many insights can be gained by analyzing such Pareto-optimal set of equations (Smits and Kotanchek, 2005; Udrescu et al., 2020; Ross et al., 2023; Beucler et al., 2024b).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

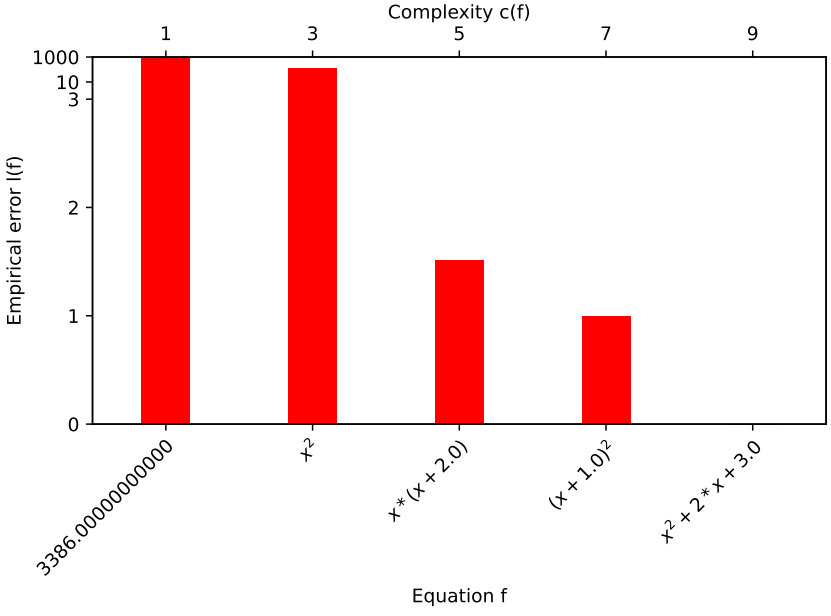

Figure A1. Example of a Pareto-optimal set of equations estimated with PySR on data generated with the equation x^2+2+3. The empirical error $l(f)$ is the mean squared error, while the complexity $c(f)$ is the total number of coefficients, operators, variables in the equation f. By definition, since this set of equations is Pareto-optimal, the error $l(f)$ decreases with the complexity $c(f)$.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Appendix B: Random sampling of hyperparameters for PySR

For each validation scheme, 500 sets of hyperparameters of PySR are sampled randomly. Each set of hyperparameters contains 29 hyperparameter values $\{\theta_1, ..., \theta_{29}\}$ that parametrize the algorithm of PySR, e.g. θ_1 is the number of iterations. For details on the 29 other chosen hyperparameters we refer to our code https://github.com/lerouxerwan/EquationDiscoveryForClimateImpact. In practice, to sample a set of hyperparameters, we sample independently 29 hyperparameter values. Specifically, if θ is the default hyperparameter value of PySR, we sample values with a log-scaled uniform distribution on the interval $[\frac{\theta}{10}, 10 \times \theta]$.

Hyperparameter	Range of values	Hyperparameter	Range of values
maxsize	[13, 40]	fraction replaced hof	[0.0061, 0.614]
warmup maxsize by	[0.1, 1.0]	weight add node	[0.247, 24.7]
populations	[2, 310]	weight insert node	[0.0011, 0.112]
population size	[17, 44]	weight delete node	[0.087, 8.7]
ncycles per iteration	[38, 3800]	weight do nothing	[0.0273, 2.73]
topn	[2, 20]	weight mutate constant	[0.0035, 0.346]
optimizer f calls limit	[1000, 100000]	weight mutate operator	[0.0293, 2.93]
optimize probability	[0.1, 1.0]	weight swap operands	[0.0198, 1.98]
tournament selection p	[0.1, 1.0]	weight rotate tree	[0.426, 42.6]
tournament selection n	[2, 20]	weight randomize	[0.0001, 0.005]
weight optimize	[0.1, 1.0]	weight simplify	[0.0002, 0.0209]
maxdepth	[2, 10]	crossover probability	[0.0026, 0.259]
niterations	[10, 1000]	perturbation factor	[0.0129, 1.29]
adaptive parsimony scaling	[104.0, 10400.0]	probability negate constant	[0.0007, 0.0743]
fraction replaced	[0.0, 0.0036]		

Table B1. The 29 hyperparameters of PySR (and their range of values) that are considered for each validation scheme (Sect. 3.3).

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Appendix C: Climate variables

Name	Abbreviation	Unit
Potential Temperature		^{o}C
Potential Density		kg/m^3
Salinity		kg/m^3
Upward Velocity	upVelo	m/s
Vertical Eddy Diffusivity		m^2/s
Zonal Current	ZonCur	m/s
Zonal Wind Stress		N/m^2
Meridional Current		m/s
Meridional Wind Stress	MWS	N/m^2
Sea Surface Temperature	SST	^{o}C
Net Downward Total Heat Flux	TH	W/m^2
Net Downward Shortwave Flux	Shortwave	W/m^2
Net Upward Water Flux		mm/s
Sea Surface Thermosteric Height		m
Sea Surface Steric Height	SSSH	m
Sea Surface Height		m
Sea Surface Salinity	SSS	kg/m^3
Surface Salt Flux		Kg/m^2/s
Mixed Layer Depth		m
Net Surface Heat Flux	NSHF	W/m^2
Baroclinic Dynamical Height		m
Surface Pressure	SP	Pa
Stratification Index	StratIndex	s^{-1}

Table C1. The 23 climate variables of CNRM-RCSM4 that are considered to design climate indicators (Sect. 2). In total we have 115 climate indicators: one for each of the 5 periods (year, winter, spring, summer, autumn) and all 23 climate variables. For visual clarity, we do not define abbreviations (the field is left empty) for variables that were not found in any of the 9 validated equations (Tab. 1)

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Appendix D: Seasonal repartition of annual net primary production

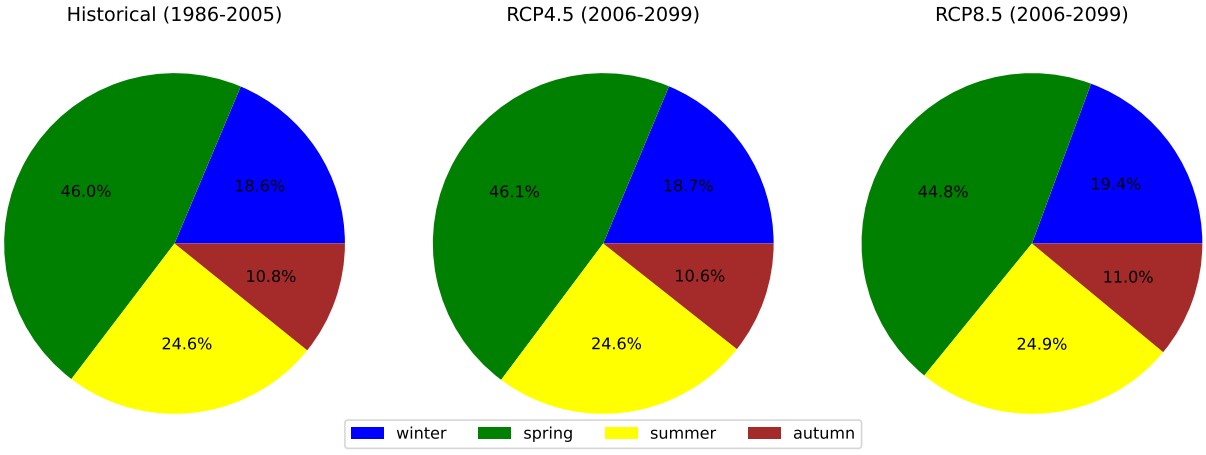

Figure D1. Seasonal repartition of annual net primary production in the studied area, i.e. an offshore area off the Gulf of Lion

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Appendix E: Four climate indicators identified from validated equations

In Figure E1, we represent the time series of the four indicators that are in a majority of the 9 equations of Table 1. Sea surface temperatures in winter (Fig. E1 **c**) and spring (Fig. E1 **b**), which inversely contributes to the predicted value, are expected to increase. By contrary, the sea surface salinity in spring (Fig. E1 **a**), that positively contributes to the predicted value, is expected to decrease from 2040. Finally, the net downward shortwave flux in winter (Fig. E1 **d**), that also contributes positively, is expected to slightly increase with an important year-to-year variability.

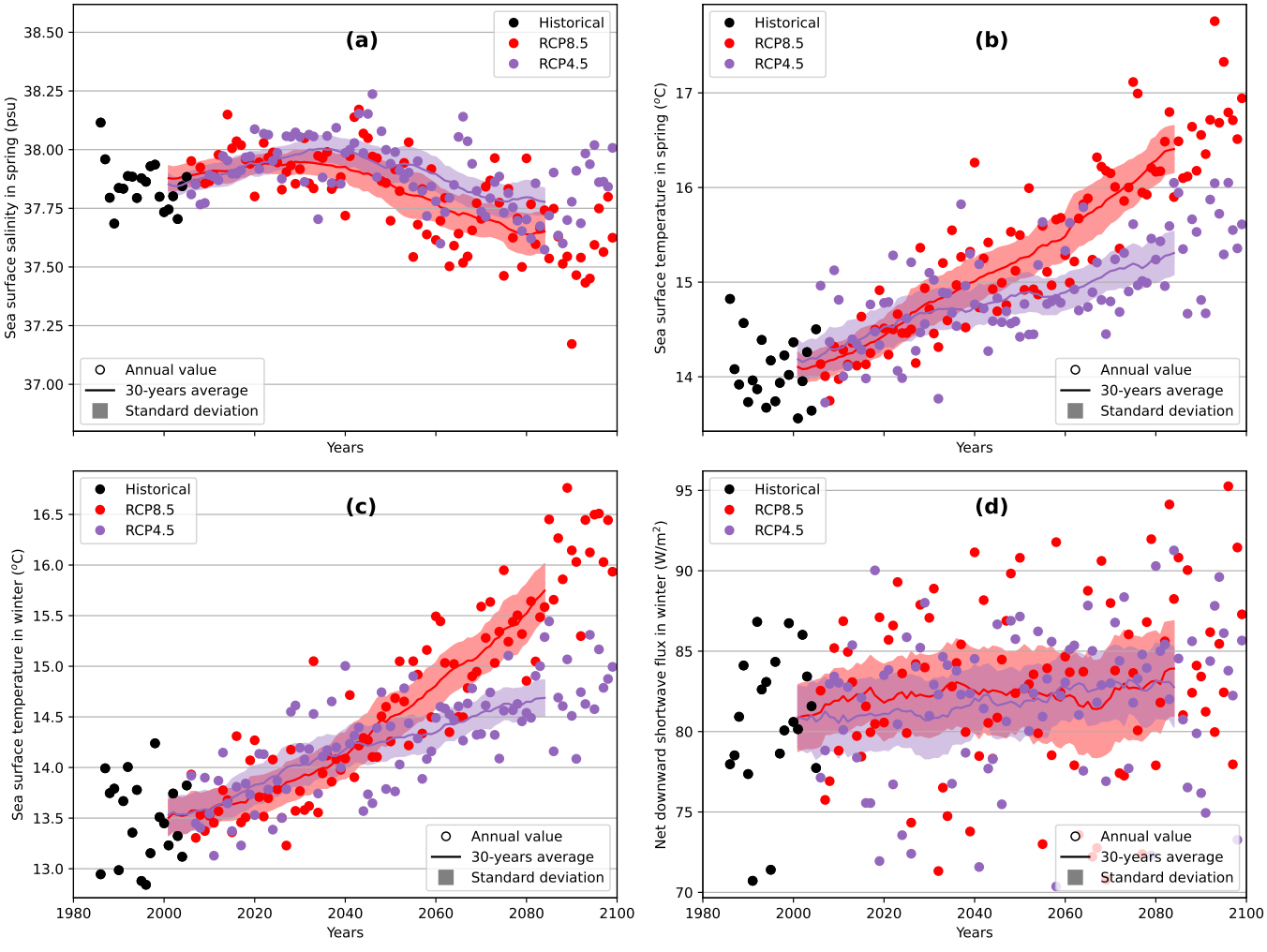

Figure E1. Time series of the four climate indicators that are in a majority of equations from Table 1. Indeed, **(a)** the sea surface salinity in spring, **(b)** the sea surface temperature in spring **(c)** the sea surface temperature in winter **(d)** the net downward shortwave flux are in 7, 6, 5, and 5 equations, respectively. Annual values (dots) are shown for the historical period and scenarios RCP4.5 and RCP8.5. For these scenarios, 30-years average (line) and standard deviation (shaded area) are also shown.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

References

Addison, H., Kendon, E., Ravuri, S., Aitchison, L., and Watson, P. A.: Machine learning emulation of precipitation from km-scale regional climate simulations using a diffusion model, http://arxiv.org/abs/2407.14158, 2024.

Atmojo, D. W., Weigel, K., Grundner, A., Holland, M. M., Sidorenko, D., and Eyring, V.: Data-driven equation discovery of a sea ice albedo parametrisation, https://doi.org/10.5194/egusphere-2025-3556, 2025.

Auger, P. A., Diaz, F., Ulses, C., Estournel, C., Neveux, J., Joux, F., Pujo-Pay, M., and Naudin, J. J.: Functioning of the planktonic ecosystem on the Gulf of Lions shelf (NW Mediterranean) during spring and its impact on the carbon deposition: A field data and 3-D modelling combined approach, Biogeosciences, 8, 3231–3261, https://doi.org/10.5194/bg-8-3231-2011, 2011.

Auger, P. A., Ulses, C., Estournel, C., Stemmann, L., Somot, S., and Diaz, F.: Interannual control of plankton communities by deep winter mixing and prey/predator interactions in the NW Mediterranean: Results from a 30-year 3D modeling study, Progress in Oceanography, 124, 12–27, https://doi.org/10.1016/j.pocean.2014.04.004, 2014.

Baklouti, M., Pagès, R., Alekseenko, E., Guyennon, A., and Grégori, G.: On the benefits of using cell quotas in addition to intracellular elemental ratios in flexible-stoichiometry Plankton functional type models. Application to the Mediterranean Sea, Progress in Oceanography, 197, https://doi.org/10.1016/j.pocean.2021.102634, 2021.

Balaji, V., Couvreux, F., Deshayes, J., Gautrais, J., Hourdin, F., and Rio, C.: Are General Circulation Models Obsolete?, 119, 1–10, www.pnas.org/cgi/doi/10.1073/pnas.XXXXXXXXXX, 2022.

Balmaseda, M., Vidard, A., and Anderson, D.: 'The NEMOVAR-COMBINE ocean re-analysis' - COMBINE Technical Report No. 1., Tech. rep., https://www.cen.uni-hamburg.de/en/icdc/data/ocean/easy-init-ocean/ecmwf-nemovar-combine.html, 2010.

Bano-Medina, J., Iturbide, M., Fernandez, J., and Gutierrez, J. M.: Transferability and explainability of deep learning emulators for regional climate model projections: Perspectives for future applications, http://arxiv.org/abs/2311.03378, 2023.

Bassetti, S., Hutchinson, B., Tebaldi, C., and Kravitz, B.: DiffESM: Conditional Emulation of Temperature and Precipitation in Earth System Models with 3D Diffusion Models, https://doi.org/10.22541/essoar.170612214.43012845/v1, 2024.

Berthold, M., Nieters, P., and Vortmeyer-Kley, R.: Machine learning to identify environmental drivers of phytoplankton blooms in the Southern Baltic Sea, Scientific Reports, 15, https://doi.org/10.1038/s41598-025-85605-y, 2025.

Beucler, T., Gentine, P., Yuval, J., Gupta, A., Peng, L., Lin, J., Yu, S., Rasp, S., Ahmed, F., O'gorman, P. A., Neelin, J. D., Lutsko, N. J., and Pritchard, M.: Climate-invariant machine learning, Tech. rep., https://www.science.org, 2024a.

Beucler, T., Grundner, A., Shamekh, S., Ukkonen, P., Chantry, M., and Lagerquist, R.: Distilling Machine Learning's Added Value: Pareto Fronts in Atmospheric Applications, http://arxiv.org/abs/2408.02161, 2024b.

Boé, J., Mass, A., and Deman, J.: A simple hybrid statistical–dynamical downscaling method for emulating regional climate models over Western Europe. Evaluation, application, and role of added value?, Climate Dynamics, 61, 271–294, https://doi.org/10.1007/s00382-022-06552-2, 2023.

Bongard, J. and Lipson, H.: Automated reverse engineering of nonlinear dynamical systems, Tech. rep., 2007.

Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., Halloran, P., Heinze, C., Ilyina, T., Séférian, R., Tjiputra, J., and Vichi, M.: Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models, Biogeosciences, 10, 6225–6245, https://doi.org/10.5194/bg-10-6225-2013, 2013.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Bosc, E., Bricaud, A., and Antoine, D.: Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations, Global Biogeochemical Cycles, 18, 10, https://doi.org/10.1029/2003GB002034ï, 2004.

Bouabid, S., Souza, A. N., and Ferrari, R.: Score-based generative emulation of impact-relevant Earth system model outputs, http://arxiv.org/abs/2510.04358, 2025.

Boussard, J., Nagda, C., Kaltenborn, J., Lange, C. E. E., Brouillard, P., Gurwicz, Y., Nowack, P., and Rolnick, D.: Towards Causal Representations of Climate Model Data, http://arxiv.org/abs/2312.02858, 2023.

Brunton, S. L., Proctor, J. L., and Kutz, J. N.: Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proceedings of the National Academy of Sciences of the United States of America, 113, 3932–3937, https://doi.org/10.1073/pnas.1517384113, 2016.

Buchanan, P. J., Reddy, P. J., Matear, R. J., Chamberlain, M. A., Rohr, T., Squire, D., and Shadwick, E. H.: Optimization of the World Ocean Model of Biogeochemistry and Trophic dynamics (WOMBAT) using surrogate machine learning methods, Biogeosciences, 22, 5349–5385, https://doi.org/10.5194/bg-22-5349-2025, 2025.

Camps-Valls, G., Gerhardus, A., Ninad, U., Varando, G., Martius, G., Balaguer-Ballester, E., Vinuesa, R., Diaz, E., Zanna, L., and Runge, J.: Discovering causal relations and equations from data, https://doi.org/10.1016/j.physrep.2023.10.005, 2023.

Castruccio, S., Mcinerney, D. J., Stein, M. L., Liu Crouch, F., Jacob, R. L., and Moyer, E. J.: Statistical Emulation of Climate Model Projections Based on Precomputed GCM Runs, https://doi.org/10.1175/JCLI-D-13, 2014.

Cornelio, C., Dash, S., Austel, V., Josephson, T. R., Goncalves, J., Clarkson, K. L., Megiddo, N., El Khadir, B., and Horesh, L.: Combining data and theory for derivable scientific discovery with AI-Descartes, Nature Communications, 14, https://doi.org/10.1038/s41467-023-37236-y, 2023.

Cozad, A. and Nikolaos, V. S.: A global MINLP approach to symbolic regression, Mathematical Programming, 170, 97–119, 2018.

Cranmer, M.: Interpretable Machine Learning for Science with PySR and SymbolicRegression.jl, http://arxiv.org/abs/2305.01582, 2023.

Darmaraki, S., Somot, S., Sevault, F., Nabat, P., Cabos Narvaez, W. D., Cavicchia, L., Djurdjevic, V., Li, L., Sannino, G., and Sein, D. V.: Future evolution of Marine Heatwaves in the Mediterranean Sea, Climate Dynamics, 53, 1371–1392, https://doi.org/10.1007/s00382-019-04661-z, 2019.

de Franca, F. O., Virgolin, M., Kommenda, M., Majumder, M. S., Cranmer, M., Espada, G., Ingelse, L., Fonseca, A., Landajuela, M., Petersen, B., Glatt, R., Mundhenk, N., Lee, C. S., Hochhalter, J. D., Randall, D. L., Kamienny, P., Zhang, H., Dick, G., Simon, A., Burlacu, B., Kasak, J., Machado, M., Wilstrup, C., and Cavaz, W. G. L.: SRBench++: Principled Benchmarking of Symbolic Regression With Domain-Expert Interpretation, IEEE Transactions on Evolutionary Computation, pp. 1–1, https://doi.org/10.1109/TEVC.2024.3423681, 2024.

D'ortenzio, F. and Ribera Dalcaì, M.: On the trophic regimes of the Mediterranean Sea: a satellite analysis, Tech. rep., www.biogeosciences.net/6/139/2009/, 2009.

Doury, A., Somot, S., Gadat, S., Ribes, A., and Corre, L.: Regional climate model emulator based on deep learning: concept and first evaluation of a novel hybrid downscaling approach, Climate Dynamics, https://doi.org/10.1007/s00382-022-06343-9, 2022.

Erlandsen, H. B., Parding, K. M., Benestad, R., Mezghani, A., and Pontoppidan, M.: A Hybrid Downscaling Approach for Future Temperature and Precipitation Change, https://doi.org/10.1175/JAMC-D-20, 2020.

Evin, G., Hingray, B., Blanchet, J., Morin, S., and Verfaillie, D.: Partitioning Uncertainty Components of an Incomplete Ensemble of Climate Projections Using Data Augmentation, https://doi.org/10.1175/JCLI-D-18, 2019.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Fabri-Ruiz, S., Berdalet, E., Ulses, C., Somot, S., Vila, M., Lemée, R., and Irisson, J. O.: Harmful Ostreopsis cf. ovata blooms could extend in time span with climate change in the Western Mediterranean Sea, Science of the Total Environment, 947, https://doi.org/10.1016/j.scitotenv.2024.174726, 2024.

Grundner, A., Beucler, T., Gentine, P., and Eyring, V.: Data-Driven Equation Discovery of a Cloud Cover Parameterization, Journal of Advances in Modeling Earth Systems, 16, https://doi.org/10.1029/2023MS003763, 2024.

Guimerà, R., Reichardt, I., Aguilar-Mogas, A., Massucci, F. A., Miranda, M., Pallarès, J., and Sales-Pardo, M.: A Bayesian machine scientist to aid in the solution of challenging scientific problems, Tech. rep., https://www.science.org, 2020.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095–1107, https://doi.org/10.1175/2009BAMS2607.1, 2009.

Hemmings, J. C., Challenor, P. G., and Yool, A.: Mechanistic site-based emulation of a global ocean biogeochemical model (MEDUSA 1.0) for parametric analysis and calibration: An application of the Marine Model Optimization Testbed (MarMOT 1.1), Geoscientific Model Development, 8, 697–731, https://doi.org/10.5194/gmd-8-697-2015, 2015.

Herrmann, M., Estournel, C., Adloff, F., and Diaz, F.: Impact of climate change on the northwestern Mediterranean Sea pelagic planktonic ecosystem and associated carbon cycle, Journal of Geophysical Research: Oceans, 119, 5815–5836, https://doi.org/10.1002/2014JC010016, 2014.

Hobeichi, S., Nishant, N., Shao, Y., Abramowitz, G., Pitman, A., Sherwood, S., Bishop, C., and Green, S.: Using Machine Learning to Cut the Cost of Dynamical Downscaling, Earth's Future, 11, https://doi.org/10.1029/2022EF003291, 2023.

Huntingford, C., Nicoll, A. J., Klein, C., and Ahmad, J. A.: Potential for equation discovery with AI in the climate sciences, Earth System Dynamics, 16, 475–495, https://doi.org/10.5194/esd-16-475-2025, 2025.

Hutchins, D. A. and Tagliabue, A.: Feedbacks between phytoplankton and nutrient cycles in a warming ocean, https://doi.org/10.1038/s41561-024-01454-w, 2024.

Kendon, E. J., Addison, H., Doury, A., Somot, S., Watson, P. A. G., Booth, B. B. B., Coppola, E., Gutiérrez, J. M., Murphy, J., and Scullion, C.: Potential for Machine Learning Emulators to Augment Regional Climate Simulations in Provision of Local Climate Change Information, Bulletin of the American Meteorological Society, 106, E1175–E1203, https://doi.org/10.1175/BAMS-D-24-0114.1, 2025.

Kommenda, M., Beham, A., Affenzeller, M., and Kronberger, G.: Complexity measures for multi-objective symbolic regression, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 9520, pp. 409–416, Springer Verlag, https://doi.org/10.1007/978-3-319-27340-2_51, 2015.

Koza, J. R.: Genetic programming as a means for programming computers by natural selection, Tech. rep., 1994.

Li, Q., Zhang, C., Wei, Z., Jin, X., Shangguan, W., Yuan, H., Zhu, J., Li, L., Liu, P., Chen, X., Yan, Y., and Dai, Y.: Advancing symbolic regression for earth science with a focus on evapotranspiration modeling, npj Climate and Atmospheric Science, 7, https://doi.org/10.1038/s41612-024-00861-5, 2024.

Macias, D. M., Garcia-Gorriz, E., and Stips, A.: Productivity changes in the Mediterranean Sea for the twenty-first century in response to changes in the regional atmospheric forcing, Frontiers in Marine Science, 2, https://doi.org/10.3389/fmars.2015.00079, 2015.

Maher, N., Milinski, S., and Ludwig, R.: Large ensemble climate model simulations: Introduction, overview, and future prospects for utilising multiple types of large ensemble, Earth System Dynamics, 12, 401–418, https://doi.org/10.5194/esd-12-401-2021, 2021.

Makke, N. and Chawla, S.: Interpretable scientific discovery with symbolic regression: a review, Artificial Intelligence Review, 57, https://doi.org/10.1007/s10462-023-10622-0, 2024.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Many, G., Ulses, C., Estournel, C., and Marsaleix, P.: Particulate organic carbon dynamics in the Gulf of Lion shelf (NW Mediterranean) using a coupled hydrodynamic-biogeochemical model, Biogeosciences, 18, 5513–5538, https://doi.org/10.5194/bg-18-5513-2021, 2021.

Martius, G. and Lampert, C. H.: Extrapolation and learning equations, http://arxiv.org/abs/1610.02995, 2016.

Mattern, J. P., Fennel, K., and Dowd, M.: Sensitivity and uncertainty analysis of model hypoxia estimates for the Texas-Louisiana shelf, Journal of Geophysical Research: Oceans, 118, 1316–1332, https://doi.org/10.1002/jgrc.20130, 2013.

McConaghy, T.: FFX: Fast, Scalable, Deterministic Symbolic Regression Technology, pp. 235–260, https://doi.org/10.1007/978-1-4614-1770-5_13, 2011.

McDermott, J., Morrow, I., Glynn, J., and Panos, E.: Symbolic Regression for Modelling Decarbonisation Pathways in the Global Energy-Economy-Climate System, in: GECCO 2025 Companion - Proceedings of the 2025 Genetic and Evolutionary Computation Conference Companion, pp. 871–874, Association for Computing Machinery, Inc, https://doi.org/10.1145/3712255.3726704, 2025.

Mora, C., Wei, C. L., Rollo, A., Amaro, T., Baco, A. R., Billett, D., Bopp, L., Chen, Q., Collier, M., Danovaro, R., Gooday, A. J., Grupe, B. M., Halloran, P. R., Ingels, J., Jones, D. O., Levin, L. A., Nakano, H., Norling, K., Ramirez-Llodra, E., Rex, M., Ruhl, H. A., Smith, C. R., Sweetman, A. K., Thurber, A. R., Tjiputra, J. F., Usseglio, P., Watling, L., Wu, T., and Yasuhara, M.: Biotic and Human Vulnerability to Projected Changes in Ocean Biogeochemistry over the 21st Century, PLoS Biology, 11, https://doi.org/10.1371/journal.pbio.1001682, 2013.

Moullec, F., Barrier, N., Drira, S., Guilhaumon, F., Marsaleix, P., Somot, S., Ulses, C., Velez, L., and Shin, Y. J.: An end-to-end model reveals losers and winners in a warming Mediterranean Sea, Frontiers in Marine Science, 6, https://doi.org/10.3389/fmars.2019.00345, 2019.

Northrop, P. J. and Chandler, R. E.: Quantifying Sources of Uncertainty in Projections of Future Climate*, https://doi.org/10.1175/JCLI-D-14, 2014.

Pagès, R., Baklouti, M., Barrier, N., Ayache, M., and Sevault, F.: Projected Effects of Climate-Induced Changes in Hydrodynamics on the Biogeochemistry of the Mediter-ranean Sea Under the RCP 8.5 Regional Climate Scenario. Frontiers in Marine Science, 7, 563 615, https://doi.org/10.3389/fmars.2020.563615ï, 2020.

Parras-Berrocal, I. M., Waldman, R., Sevault, F., Somot, S., Gonzalez, N., Ahrens, B., Anav, A., Djurdjević, V., Gualdi, S., Hamouda, M. E., Li, L., Lionello, P., Sannino, G., and Sein, D. V.: Response of the Mediterranean Sea Surface Circulation at Various Global Warming Levels: A Multi-Model Approach, Geophysical Research Letters, 51, https://doi.org/10.1029/2024GL111695, 2024.

Petersen, B. K., Landajuela, M., Mundhenk, T. N., Santiago, C. P., Kim, S. K., and Kim, J. T.: Deep symbolic regression: Recovering mathematical expressions from data via risk-seeking policy gradients, http://arxiv.org/abs/1912.04871, 2019.

Pirani, A., Fuglestvedt, J. S., Byers, E., O'Neill, B., Riahi, K., Lee, J.-Y., Marotzke, J., Rose, S. K., Schaeffer, R., and Tebaldi, C.: Scenarios in IPCC assessments: lessons from AR6 and opportunities for AR7, npj Climate Action, 3, 1, https://doi.org/10.1038/s44168-023-00082-1, 2024.

Rampal, N., Gibson, P. B., Sherwood, S., and Abramowitz, G.: On the Extrapolation of Generative Adversarial Networks for Downscaling Precipitation Extremes in Warmer Climates, Geophysical Research Letters, 51, https://doi.org/10.1029/2024GL112492, 2024.

Richon, C., Dutay, J. C., Bopp, L., Le Vu, B., Orr, J. C., Somot, S., and Dulac, F.: Biogeochemical response of the Mediterranean Sea to the transient SRES-A2 climate change scenario, Biogeosciences, 16, 135–165, https://doi.org/10.5194/bg-16-135-2019, 2019.

Ross, A., Li, Z., Perezhogin, P., Fernandez-Granda, C., and Zanna, L.: Benchmarking of Machine Learning Ocean Subgrid Parameterizations in an Idealized Model, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022MS003258, 2023.

Ruti, P. M., Somot, S., Giorgi, F., Dubois, C., Flaounas, E., Obermann, A., Dell'Aquila, A., Pisacane, G., Harzallah, A., Lombardi, E., Ahrens, B., Akhtar, N., Alias, A., Arsouze, T., Aznar, R., Bastin, S., Bartholy, J., Béranger, K., Beuvier, J., Bouffies-Cloché, S., Brauch, J.,

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

Cabos, W., Calmanti, S., Calvet, J. C., Carillo, A., Conte, D., Coppola, E., Djurdjevic, V., Drobinski, P., Elizalde-Arellano, A., Gaertner, M., Galàn, P., Gallardo, C., Gualdi, S., Goncalves, M., Jorba, O., Jordà, G., L'Heveder, B., Lebeaupin-Brossier, C., Li, L., Liguori, G., Lionello, P., Maciàs, D., Nabat, P., Önol, B., Raikovic, B., Ramage, K., Sevault, F., Sannino, G., Struglia, M. V., Sanna, A., Torma, C., and Vervatis, V.: Med-CORDEX initiative for Mediterranean climate studies, Bulletin of the American Meteorological Society, 97, 1187–1208, https://doi.org/10.1175/BAMS-D-14-00176.1, 2016.

Sahoo, S. S., Lampert, C. H., and Martius, G.: Learning Equations for Extrapolation and Control, Tech. rep., 2018.

Santer, B., Wigley, T., Schlesinger, M., and Mitchell, J.: Developing Climate Scenarios From Equilibrium, Tech. rep., 1990.

Schartau, M., Wallhead, P., Hemmings, J., Löptien, U., Kriest, I., Krishna, S., Ward, B. A., Slawig, T., and Oschlies, A.: Reviews and syntheses: Parameter identification in marine planktonic ecosystem modelling, Biogeosciences, 14, 1647–1701, https://doi.org/10.5194/bg-14-1647-2017, 2017.

Schmidt, M. and Lipson, H.: Distilling Free-Form Natural Laws from Experimental Data, Tech. rep., www.sciencemag.org, 2009.

Sevault, F., Somot, S., Alias, A., Dubois, C., Lebeaupin-Brossier, C., Nabat, P., Adloff, F., Déqué, M., and Decharme, B.: A fully coupled Mediterranean regional climate system model: design and evaluation of the ocean component for the 1980–2012 period, Tellus A: Dynamic Meteorology and Oceanography, 66, 23 967, https://doi.org/10.3402/tellusa.v66.23967, 2014.

Sigman, D. M. and Hain, M. P.: The Biological Productivity of the Ocean, Tech. rep., http://bats.bios.edu, 2012.

Skákala, J., Awty-Carroll, K., Menon, P. P., Wang, K., and Lessin, G.: Future digital twins: emulating a highly complex marine biogeochemical model with machine learning to predict hypoxia, Frontiers in Marine Science, 10, https://doi.org/10.3389/fmars.2023.1058837, 2023.

Smits, G. F. and Kotanchek, M.: Pareto-Front Exploitation in Symbolic Regression, in: Genetic programming theory and practice II, chap. 17, 2005.

Somot, S., Sevault, F., and Déqué, M.: Transient climate change scenario simulation of the Mediterranean Sea for the twenty-first century using a high-resolution ocean circulation model, Climate Dynamics, 27, 851–879, https://doi.org/10.1007/s00382-006-0167-z, 2006.

Song, W., Jiang, S., Camps-Valls, G., Williams, M., Zhang, L., Reichstein, M., Vereecken, H., He, L., Hu, X., and Shi, L.: Towards data-driven discovery of governing equations in geosciences, Communications Earth & Environment, 5, 589, https://doi.org/10.1038/s43247-024-01760-6, 2024.

Soto-Navarro, J., Jordá, G., Amores, A., Cabos, W., Somot, S., Sevault, F., Macías, D., Djurdjevic, V., Sannino, G., Li, L., and Sein, D.: Evolution of Mediterranean Sea water properties under climate change scenarios in the Med-CORDEX ensemble, Climate Dynamics, 54, 2135–2165, https://doi.org/10.1007/s00382-019-05105-4, 2020.

Stanislawska, K., Krawiec, K., and Kundzewicz, Z. W.: Modeling global temperature changes with genetic programming, Computers and Mathematics with Applications, 64, 3717–3728, https://doi.org/10.1016/j.camwa.2012.02.049, 2012.

Takakura, J., Fujimori, S., Takahashi, K., Hanasaki, N., Hasegawa, T., Hirabayashi, Y., Honda, Y., Iizumi, T., Park, C., Tamura, M., and Hijioka, Y.: Reproducing complex simulations of economic impacts of climate change with lower-cost emulators, Geoscientific Model Development, 14, 3121–3140, https://doi.org/10.5194/gmd-14-3121-2021, 2021.

Tebaldi, C., Snyder, A., and Dorheim, K.: STITCHES: creating new scenarios of climate model output by stitching together pieces of existing simulations, Earth System Dynamics, 13, 1557–1609, https://doi.org/10.5194/esd-13-1557-2022, 2022.

Udrescu, S.-M. and Tegmark, M.: AI Feynman: a Physics-Inspired Method for Symbolic Regression, http://arxiv.org/abs/1905.11481, 2019.

Udrescu, S.-M., Tan, A., Feng, J., Neto, O., Wu, T., and Tegmark, M.: AI Feynman 2.0: Pareto-optimal symbolic regression exploiting graph modularity, Tech. rep., https://ai-feynman., 2020.

https://doi.org/10.5194/egusphere-2026-1991
Preprint. Discussion started: 12 May 2026

van der Meer, M., de Roda Husman, S., and Lhermitte, S.: Deep Learning Regional Climate Model Emulators: A Comparison of Two Downscaling Training Frameworks, Journal of Advances in Modeling Earth Systems, 15, https://doi.org/10.1029/2022ms003593, 2023.

Virgolin, M. and Pissis, S. P.: Symbolic Regression is NP-hard, http://arxiv.org/abs/2207.01018, 2022.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M. P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: Description and basic evaluation, Climate Dynamics, 40, 2091–2121, https://doi.org/10.1007/s00382-011-1259-y, 2013.

Watson-Parris, D., Rao, Y., Olivié, D., Seland, Nowack, P., Camps-Valls, G., Stier, P., Bouabid, S., Dewey, M., Fons, E., Gonzalez, J., Harder, P., Jeggle, K., Lenhardt, J., Manshausen, P., Novitasari, M., Ricard, L., and Roesch, C.: ClimateBench v1.0: A Benchmark for Data-Driven Climate Projections, Journal of Advances in Modeling Earth Systems, 14, https://doi.org/10.1029/2021MS002954, 2022.

Wu, T. and Tegmark, M.: Toward an AI Physicist for Unsupervised Learning, https://doi.org/10.1103/PhysRevE.100.033311, 2018.

Yu, Z., Wang, G., Ding, J., Wang, H., and Li, Y.: Beyond Formula Complexity: Effective Information Criterion Improves Performance and Interpretability for Symbolic Regression, http://arxiv.org/abs/2509.21780, 2025.