

# Improving statistical prediction and revealing nonlinearity of ENSO using observations of ocean heat content in the tropical Pacific

Aleksei Seleznev (✉ [aseleznev@ipfran.ru](mailto:aseleznev@ipfran.ru))

Institute of Applied Physics RAS: FGBNU Federal'nyj issledovatel'skij centr Institut prikladnoj fiziki  
Rossijskoj akademii nauk <https://orcid.org/0000-0002-6321-8109>

Dmitry Mukhin

Institute of Applied Physics RAS: FGBNU Federal'nyj issledovatel'skij centr Institut prikladnoj fiziki  
Rossijskoj akademii nauk

---

## Research Article

**Keywords:** statistical ENSO models, data-driven models, spring predictability barrier, early predictors of ENSO, ENSO nonlinearity, ocean heat content

**Posted Date:** December 29th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1207911/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

---

1 **Improving statistical prediction and revealing**  
2 **nonlinearity of ENSO using observations of ocean**  
3 **heat content in the tropical Pacific**

4 **Aleksei Seleznev · Dmirty Mukhin**

5  
6 Received: date / Accepted: date

7 **Abstract** It is well-known that the upper ocean heat content (OHC) variabil-  
8 ity in the tropical Pacific contains valuable information about dynamics of El  
9 Niño–Southern Oscillation (ENSO). Here we combine sea surface temperature  
10 (SST) and OHC indices derived from the gridded datasets to construct a phase  
11 space for data-driven ENSO models. Using a Bayesian optimization method,  
12 we construct linear as well as nonlinear models for these indices. We find that  
13 the joint SST-OHC optimal models yield significant benefits in predicting both  
14 the SST and OHC as compared with the separate SST or OHC models. It is  
15 shown that these models substantially reduces seasonal predictability barriers  
16 in each variable – the spring barrier in the SST index and the winter barrier in  
17 the OHC index. We also reveal the significant nonlinear relationships between  
18 the ENSO variables manifesting on interannual scales, which opens prospects  
19 for improving yearly ENSO forecasting.

20 **Keywords** statistical ENSO models · data-driven models · spring pre-  
21 dictability barrier · early predictors of ENSO · ENSO nonlinearity · ocean  
22 heat content

23 **1 Introduction**

24 El Niño–Southern Oscillation (ENSO) is the dominant mode of interannual cli-  
25 mate variability which originates in the tropical Pacific, but impacts climate  
26 conditions over the world (Trenberth, 2019; Alexander et al., 2002; Wang and  
27 Picaut, 2004). Historically, two conceptual elements are considered as key in-  
28 gredients underlying ENSO. The first one is a Bjerknes mechanism (Bjerknes,

---

A. Seleznev, D. Mukhin  
Institute of Applied Physics of RAS, 46 Ul'yanov Str., 603950 Nizhny Novgorod, Russia  
Tel.: +7-831-4189044  
Fax: +7-831-4160616  
E-mail: aseleznev@ipfran.ru

1969) based on positive ocean-atmosphere feedback: weakening of the trade winds in response to increasing sea surface temperature (SST) results in even warmer SST in the equatorial eastern and central Pacific. The second was realized by Wyrтки (1975, 1985), who supposed that accumulation of warm water in the equatorial Pacific is a necessary precondition for the initiation of a warm ENSO event (El Niño). Strong trade winds contribute to accumulating warm water in the western part of the basin, thus building up of the east-west slope of sea level. Eventually, excessive amount of warm water provides favorable conditions for triggering the Bjerknes feedback yielding the weakening of the trade winds due to increasing of SST that contributes to eastward transport of accumulated warm water. Further studies developed the Bjerkens-Wyrтки hypothesis to explain the distinctive cyclic nature of ENSO. In the so-called recharging oscillator theory of ENSO, charge-discharge of the warm water, and hence, the heat content in the tropical Pacific is regarded as a key process underlying the observed oscillations (Cane and Zebiak, 1985; Jin, 1997). This theory involves the meridional subsurface water transport driven by the wind stress curl (also known as Sverdrup transport) as the main source of the heat content alteration. The anomalous heat content stored in the tropics due to the equatorward mass transport during the cold (La Niña) and neutral phases of ENSO eventually enables an El Niño event onset, which, in turn, changes the wind stress curl outside the equator, and, as a result, discharges warm water. After that the charging stage of the oscillation starts again. Alternative theory of ENSO is based on the delayed oscillator models (Suarez and Schopf, 1988; Galanti and Tziperman, 2000) highlighting the role of oceanic equatorial waves as carriers of thermocline depth anomalies along the equator. Such anomalies impact the SST and therefore can lead to initiating the Bjerknes feedback. Different directions and propagating times inherent for different equatorial wave modes provide complex quasiperiodical ENSO-like oscillations in such models. All the key physical processes both theories account for are shown take place in the coupled shallow-water ocean-atmosphere models (Zebiak and Cane, 1987; Anderson and McCreary, 1985; Jin and Neelin, 1993). The role of stochastic forcing in ENSO dynamics is also important, as was noticed, e.g., in (Philander and Fedorov, 2003; Fedorov et al., 2003; Chen et al., 2016; Hu and Fedorov, 2019; Martinez-Villalobos et al., 2019), since it is responsible for ENSO irregularity. Typically, it is associated with an atmospheric noise producing short-scale zonal wind anomalies (e.g., westerly wind bursts (Levine and Jin, 2017; Hu and Fedorov, 2019)). Among the drivers of such anomalies are indicated, for example, the Madden–Julian oscillation (Zhang and Gottschalck, 2002; Chiodi et al., 2014; Puy et al., 2016), or large-scale subtropical atmospheric patterns (Vimont et al., 2003; Sullivan et al., 2021).

Growing amount of high-resolution measurements of different geophysical fields in recent decades offers great opportunities of verifying existing concepts of ENSO as well as for constructing data-driven prognostic models. The data-driven, or statistical, ENSO models became an efficient tool for interseasonal ENSO forecasting; they can compete with dynamical, i.e. constructed from the “first principles”, models in this regard (Barnston et al., 2012). The com-

75 mon problem for both statistical and dynamical ENSO models is the spring  
76 predictability barrier (SPB) (Jin et al., 2008; Barnston et al., 2012) which  
77 substantially limits the tropical SST forecasts that start from the winter and  
78 spring seasons. Many statistical ENSO models (Penland and Sardeshmukh,  
79 1995; Kondrashov et al., 2005; Gavrilo et al., 2019) are based on purely SST  
80 anomalies in the tropical Pacific which accurate forecast is the main goal in  
81 ENSO predictive modeling (Barnston et al., 2012). In such models, the SPB is  
82 caused by the observed loss of autocorrelations in tropical SST with a peak in  
83 May – June. Relying on theoretical understanding of ENSO, many studies are  
84 focused on finding additional atmospheric and oceanic predictors which can  
85 help to lower the SPB. Various predictors based on ocean heat content (OHC)  
86 (Clarke and Van Gorder, 2003), warm water volume (Meinen and McPhaden,  
87 2000; Chen et al., 2020), as well as atmospheric fields (Clarke and Van Gorder,  
88 2003; Byshev et al., 2016; Chen et al., 2020; Mukhin et al., 2021) has been sug-  
89 gested. Nevertheless, there is still no conventional way to derive statistically  
90 justified predictors from data and to include them into prognostic models. Of-  
91 ten (Chen et al., 2020; Mukhin et al., 2021) such predictors are determined by  
92 finding significant lagged correlations between time series of SST-based ENSO  
93 index which needs to be predicted and corresponding time series of another  
94 ENSO-related climate variables. Typically, the obtained predictors are passed  
95 to the model as a fixed forcing (e.g. as components of regression (Clarke and  
96 Van Gorder, 2003; Chen et al., 2020)), but not as dynamical variables, which  
97 makes it difficult to use such models for “no look ahead” forecast requiring  
98 extrapolation of the predictors to the future.

99 In this study we introduce an efficient predictor of ENSO-related SST  
100 variability constructed from OHC anomalies in the tropical Pacific. The pro-  
101 posed signal is obtained simply using the standard empirical orthogonal func-  
102 tion (EOF) decomposition. We construct an optimal data-driven ENSO model  
103 which uses this predictor along with the SST-based predictor as equitable dy-  
104 namical variables. Being phase-shifted, these SST- and OHC-based variables  
105 complement each other providing proper phase space capturing the ENSO dy-  
106 namics. We demonstrate that the model obtained surpasses the purely SST-  
107 based model in predicting SST variability and allows to substantially lower the  
108 SPB. Also, we show that joint analysis of the SST- and OHC-based variables  
109 uncovers the long-term nonlinear relationships between the ENSO variables,  
110 thus revealing ENSO nonlinearity on interannual time scales.

111 The paper is organized as follows. In Sec. 2 we present a general form of the  
112 proposed data-driven model of ENSO, outline its phase space, parameteriza-  
113 tions and the learning procedure. In Sec. 3, we describe the analyzed data and  
114 the EOF analysis used for obtaining the variables capturing the meaningful  
115 processes contributing to ENSO dynamics. The different data-driven stochas-  
116 tic models (linear and nonlinear) based on obtained variables are compared.  
117 Then we analyze prediction skills and qualitative properties of the models. In  
118 Sec. 4 we discuss the obtained results and conclude.

## 2 Data-driven ENSO model

### 2.1 Phase space of the model

In constructing our ENSO model we use the concept of data-driven stochastic model developed in (Molkov et al., 2012; Mukhin et al., 2015b; Gavrillov et al., 2017) and adapted for high-dimensional and spatially distributed data in (Mukhin et al., 2015a; Gavrillov et al., 2019). Let the time series  $\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_N)$ ,  $\mathbf{x}_n \in \mathbb{R}^D$  represents observations of some ENSO-related climate variable obtained in  $D$  nodes of a spatial grid at equidistant time moments  $t_1, \dots, t_N$ . Without loss of generality, we suppose that the time series is monthly sampled and has zero mean, i.e.  $\frac{1}{N} \sum_{n=1}^N \mathbf{x}_n = \mathbf{0}$ . We use the conventional Empirical orthogonal function (EOF) analysis (Hannachi et al., 2007) to construct the phase space of the ENSO model from observed data  $\mathbf{X}$ . The corresponding state variables are obtained as  $d$  leading principal components (PCs)  $\mathbf{p}_n = \mathbf{V}^T \mathbf{x}_n$ ,  $\mathbf{p}_n \in \mathbb{R}^d$ , i.e. the projections of data vectors  $\mathbf{x}_n$  at time  $t_n$  to  $d$  EOFs (columns of the  $D \times d$  matrix  $\mathbf{V}$ ), that explain a substantial part of data variance:  $\sum_{k=1}^d \langle p_{k,n}^2 \rangle_n$ . The transformation from PCs space back into physical space is a linear map:

$$\mathbf{x}_n = \mathbf{V} \mathbf{p}_n + \mathbf{V}' \mathbf{p}'_n, \quad (1)$$

where  $\mathbf{V}'$  is a  $D \times D - d$  matrix, which columns are the residual EOFs and  $\mathbf{p}'_n \in \mathbb{R}^{D-d}$  are the corresponding PCs.

Although the leading EOFs characterize the most meaningful processes contributing to the observed dynamics, the residual EOFs keep the useful information about short autocorrelations in the observed dynamics, which could improve the short-term prediction of a state trajectory. In this work we construct the evolution model for the leading and residual PCs separately. The particular functional form of the corresponding models in the context of ENSO modeling is described in the next section.

### 2.2 Functional form of the model

#### 2.2.1 Leading PCs

The general form of the model we use for describing evolution of leading PCs is a stochastic model with memory (Molkov et al., 2012; Mukhin et al., 2015a; Gavrillov et al., 2017):

$$\mathbf{p}_n = \mathbf{f}(\mathbf{p}_{n-1}, \dots, \mathbf{p}_{n-l}) + \hat{\mathbf{g}} \cdot \boldsymbol{\xi}_n. \quad (2)$$

Here the first term  $\mathbf{f}$  is a deterministic function depending on  $l$  successive states of the system. The second term in (2) is a random component aimed at

152 modeling poorly resolved processes (e.g., the processes which time scales are  
 153 close to the sampling time). This component is expressed as the product of a  
 154 low-triangular deterministic  $d \times d$  matrix  $\widehat{\mathbf{g}}$  and a random vector  $\boldsymbol{\xi}_n \in \mathbb{R}^d$  which  
 155 is assumed taken from Gaussian uncorrelated (in space and time) processes  
 156 with zero means and unit variances. Resulting noise in the model has the  
 157 covariance matrix  $\widehat{\mathbf{g}}\widehat{\mathbf{g}}^T$ . Note that neither parameters of the function  $\mathbf{f}$  nor the  
 158 matrix  $\widehat{\mathbf{g}}$  are known *a priori*; they need to be estimated through model learning.

159 In this work we use two different parameterizations of deterministic part  $\mathbf{f}$   
 160 of the model (2) which account phase locking of the ENSO dynamics to the  
 161 annual cycle (Chen and Jin, 2020). The first one is a linear parameterization,  
 162 suggested by Mukhin et al. (2021):

$$\mathbf{f}(\mathbf{z}_n) = \mathbf{A}_n \mathbf{z}_n. \quad (3)$$

163 Here  $\mathbf{z}_n \in \mathbb{R}^{ld}$  contains the components of the vectors  $\mathbf{p}_{n-1}, \dots, \mathbf{p}_{n-l}$ ,  $\mathbf{A}_n$   
 164 is a  $d \times ld$  matrix of coefficients. To model the seasonal forcing needed for  
 165 accounting possible annual cycles in data, the coefficients are defined to be  
 166 periodic with the period  $T = 12$  month. They are decomposed into the discrete  
 167 Fourier series:

$$\mathbf{A}_n = \mathbf{A}^0 + \sum_{k=1}^q \left[ \mathbf{A}_c^k \cos \frac{2\pi k}{T} n + \mathbf{A}_s^k \sin \frac{2\pi k}{T} n \right], \quad (4)$$

168 where the parameter  $q$  taking values from 0 to 6 ( $\mathbf{A}_s^6 = \mathbf{0}$  by definition; the case  
 169  $q = 0$  corresponds to a simple linear model with constant  $\mathbf{A}_n = \mathbf{A}^0$ ) regulates  
 170 possible dependence of the model on different harmonics of the annual cycle.

171 The second parameterization we consider is nonlinear. In this case the  
 172 deterministic part  $\mathbf{f}$  of the model is represented by a single layer perceptron  
 173 with the hyperbolic tangent activation function:

$$\mathbf{f}(\mathbf{z}_n) = \sum_{i=1}^m \boldsymbol{\alpha}_i \tanh(\boldsymbol{\omega}_i^T \mathbf{z}_n + \boldsymbol{\delta}_i^T \mathbf{c}_n + \gamma_i). \quad (5)$$

174 Here  $\mathbf{c}_n = (\cos \frac{2\pi}{T} n, \sin \frac{2\pi}{T} n)$  is a two-dimensional harmonic signal which is  
 175 passed to the model input together with the state vector  $\mathbf{z}_n$  in order to model  
 176 the seasonal forcing,  $T = 12$  month,  $\boldsymbol{\alpha}_i \in \mathbb{R}^d$ ,  $\boldsymbol{\omega}_i \in \mathbb{R}^{ld}$ ,  $\boldsymbol{\delta}_i \in \mathbb{R}^2$ ,  $\gamma_i \in \mathbb{R}$  are  
 177 the unknown coefficients. The function in the form (5) is able to approximate  
 178 an arbitrary nonlinear dependence just by increasing the number of neurons  $m$   
 179 (Cybenko, 1989). The efficiency of such a parameterization in different ENSO-  
 180 related examples was demonstrated in (Mukhin et al., 2015b,a; Gavrilo et al.,  
 181 2019).

182 Given some fixed value of the leading PCs  $d$ , the complexity of the model  
 183 deterministic part  $\mathbf{f}$  is defined by its structural parameters (or hyperparameters)  
 184  $l, q$  in the case of linear parameterization (3)-(4) and  $l, m$  in the case of non-  
 185 linear parameterization (5). To avoid overfitting of the model, the choice of  
 186 the hyperparameters should be statistically justified, or optimal. According to

187 (Gavrilov et al., 2017, 2019; Mukhin et al., 2021), we use the Bayesian opti-  
 188 mality criterion for estimating them, which relies on assessing the probability  
 189 density function of data given the particular model; see details in Appendix  
 190 A.

### 191 2.2.2 Residual PCs

192 We construct the model for the residual PCs in the same way as described  
 193 in (Gavrilov et al., 2019). According to this work, the evolution law of each  
 194  $k$ -th residual PC  $\{p'_{k,n}\}$  ( $k = 1, \dots, D - d$ ) is approximated by the first-order  
 195 autoregressive model separately:

$$p'_{k,n} = b_k p'_{k,n-1} + \sigma_k \cdot \eta_{k,n}. \quad (6)$$

196 Here  $\{\eta_{k,n}\}$  is a sample of the uncorrelated Gaussian noise with the variance  
 197 equal to 1 and zero mean,  $b_k$  and  $\sigma_k$  are the parameters estimated by the least  
 198 square method. In doing so, we represent the residual PCs as independent red  
 199 noise processes. Including such a model in the forecasting scheme is aimed at  
 200 improving the prediction skills at lead times of order of autocorrelation times  
 201 of the processes captured by the residual PCs.

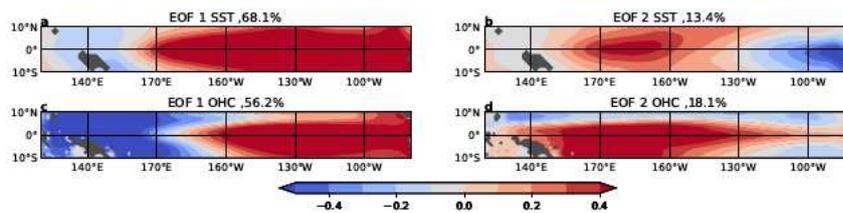
## 202 3 Results

### 203 3.1 Data and preprocessing

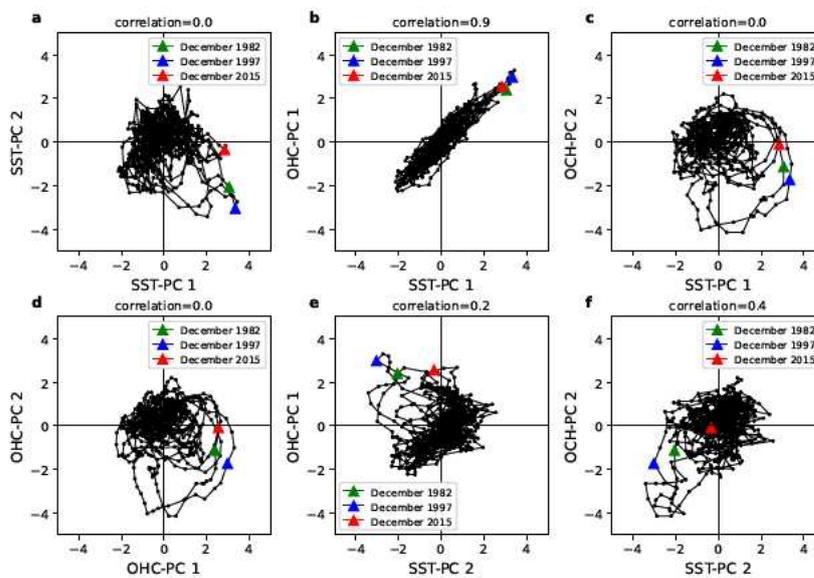
204 We construct the data-driven model from two datasets reflecting ENSO-related  
 205 variability. The first one is the monthly sea surface temperature (SST) taken  
 206 from Extended Reconstructed SST (ERSST) data set (version 5) with  $2^0 \times 2^0$   
 207 spatial resolution (Huang et al., 2017). The second dataset is the monthly time  
 208 series of ocean heat content (OHC) in 0-300 m depth layer defined on a  $1^0 \times 1^0$   
 209 grid provided by the Institute of Atmospheric Physics (Cheng et al., 2017).  
 210 From both datasets we took data in the tropical Pacific region (10S-10N, 120E-  
 211 80W) covering the time interval from Jan 1960 to Dec 2020; the total duration  
 212 of the time series is  $N = 732$  months. The anomalies were prepared from this  
 213 data by subtracting the monthly climatology within the 1960–2020 interval  
 214 followed by removing the linear regression on the  $\text{CO}_2$  trend.

### 215 3.2 EOF analysis

216 Figure 1 shows the spatial patterns corresponding to the two leading EOFs of  
 217 the sea surface temperature anomalies (SSTA) and ocean heat content anoma-  
 218 lies (OHCA) fields obtained as described in Sec. 3.1. For both data sets they  
 219 explain more than 70 % of data variance. It is often noted (Martinez-Villalobos  
 220 et al., 2019; Deser et al., 2009; Bamston et al., 1997) that the first EOF of  
 221 SSTA in the tropical Pacific is associated with ENSO and the corresponding



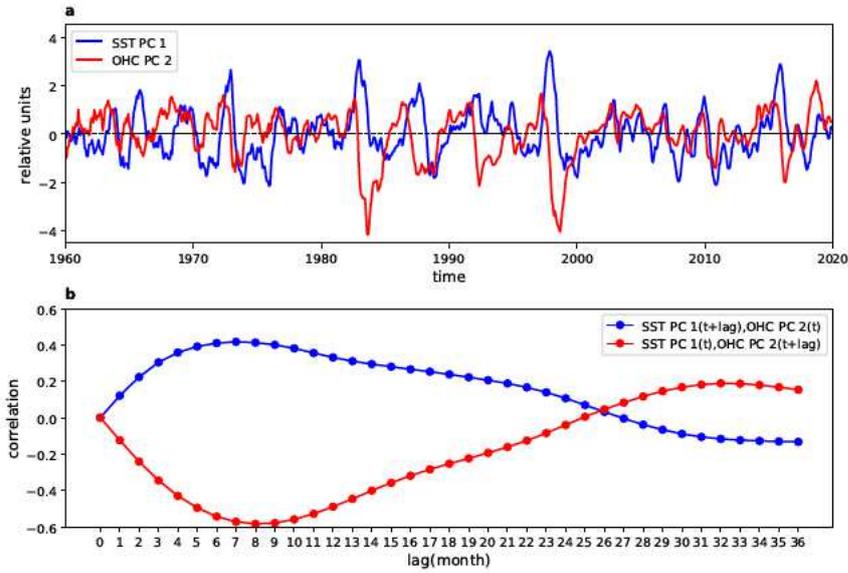
**Fig. 1** Spatial patterns corresponding to the two leading EOFs of SSTA (top panel, color scale in  $^{\circ}C$ ) and OHCA (bottom panel, color scale in  $10^9 J/m^2$ ) fields. Fractions of explained variance in percentage are shown for each EOF in the titles of the panels.



**Fig. 2** Phase trajectories reflecting the relationships between the leading PCs of SSTA and OHCA in different planes. The color markers correspond to December of the year of a very strong El Niño event (as classified by <https://ggweather.com/enso/oni.htm>). All quantities are normalized by their standard deviations. The absolute values of the Pearson correlation coefficient between the PCs are shown in the titles of the panels.

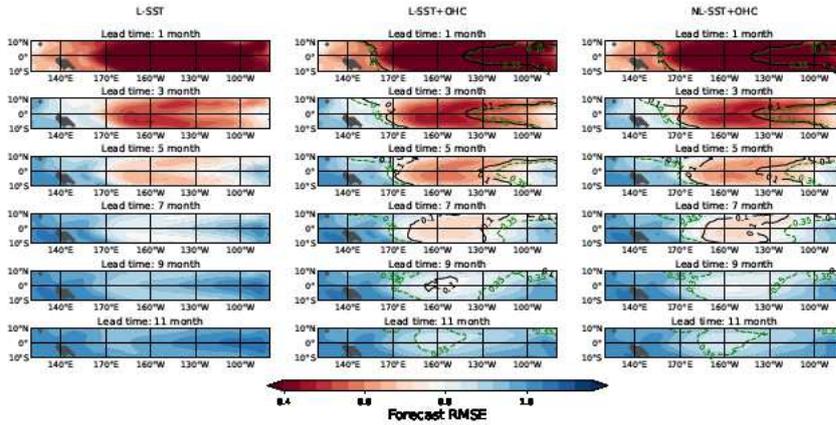
222 PC strongly correlates with the Niño 3.4 index. The second EOF together with  
 223 the first EOF allow to describe the diversity of ENSO, i.e. variety of SSTA  
 224 patterns arising during different El Niño events (e.g., “canonical” or “Modoki”  
 225 El Niño (Takahashi et al., 2011)).

226 In order to interpret the OHCA EOFs we consider the planes of different  
 227 combinations of the two leading SSTA and OHCA PCs (Fig. 2). It is clearly



**Fig. 3** Comparison of the SSTA PC1 and the OHCA PC2. Top panel: time series of the SSTA PC1 (blue) and the OHCA PC2 (red), divided by their standard deviations. Bottom panel: lagged cross correlations between these time series (see the legend).

228 seen from this figure that the first OHCA and SSTA PCs strongly correlate  
 229 (Fig. 2 (b)). Note that the corresponding EOF patterns shown in Fig. 1 (a),(c)  
 230 are also similar. What we can learn from the Fig. 2 (c)-(d) is a cyclic nature  
 231 of trajectories in both SSTA PC1-OHCA PC2 and OHCA PC1-PC2 planes  
 232 indicating an apparent phase shift between these variables. We obtain that  
 233 the peak absolute value of correlation between the SSTA PC1 and OHCA  
 234 PC2 is achieved with a lag of about 5-9 months, as Fig. 3 (b) demonstrates.  
 235 We note that the similar results about the relationships between SST and  
 236 OHC in the tropical Pacific have been obtained by Meinen and McPhaden  
 237 (2000) and Clarke et al. (2007) using warm water volume observations. The  
 238 EOF pattern corresponding to the second OHCA PC (Fig. 1 (d)) dominates  
 239 mainly in the central and western tropical Pacific and can be associated with  
 240 the OHC accumulation and discharge before and during the El Niño events  
 241 (Zebiak, 1989; Clarke et al., 2007; Cheng et al., 2019).



**Fig. 4** Spatial distributions of the SSTA forecast RMSE (normalized by the standard deviation of the  $CO_2$  detrended SSTA time series at each grid point) given by three different models for different lead time. The contours in the central and right panels bound the areas of significant improvements for the forecast RMSE of joint L-SST+OHC and NL-SST+OHC models over the separate L-SST model.

### 242 3.3 ENSO modeling

243 In this section we analyze prediction skills and qualitative properties of the  
 244 different data-driven ENSO models, built in accordance with the scheme de-  
 245 scribed in the section 2. As it follows from the analysis performed in the  
 246 previous section, the first SSTA and second OHCA PCs contain most useful  
 247 information about ENSO dynamics. According to the concept developed in  
 248 section 2.1 we treat them as the state variables of the stochastic model with  
 249 memory of the form (2). Here we construct and compare the stochastic models  
 250 of the following types:

- 251 1. Separate linear models (3)-(4), the first learnt from the SSTA PC1 (L-SST  
 252 model) and the second – from the OHCA PC2 (L-OHC model);
- 253 2. Joint linear model (3)-(4), i. e. the model learnt from both these PCs (L-  
 254 SST+OHC model);
- 255 3. Joint nonlinear model (5) (NL-SST+OHC model).

256 We train each model using the Bayesian optimization procedure described in  
 257 Appendix A. The estimated optimal values of the hyperparameters are  $l = 2$   
 258 and  $q = 1$  for the L-SST and L-SST+OHC models,  $l = 3$  and  $q = 1$  for the  
 259 L-OHC model and  $l = 2$  and  $m = 5$  for the NL-SST+OHC model. Note that  
 260 the the obtained  $q = 1$  implies the coefficients of a model are the sine functions  
 261 with the period  $T = 12$  month, see Eq. (4). The future states of the system  
 262 can be predicted by iterating these models starting from the current states.  
 263 Then the predicted values of PCs can be transformed into the physical space  
 264 using Eq. (1). Since the described models are stochastic (see Eqs. (2) and (6)),

the forecasts they produce are random sequences of states, which means that different model runs yield different forecasts. According with (Gavrilov et al., 2019; Mukhin et al., 2021), the future value of a quantity  $x$  is assessed from a model forecast as the ensemble median  $\bar{x}$  over a large number of the model runs.

### 3.3.1 Prediction skill analysis

To analyze and compare prediction skills of the data-driven models, we use two conventional metrics. The first one is the root mean square forecast error (RMSE), defined through the differences between the true and predicted values of a variable of interest  $x$ , for the time instances inside the learning set:

$$e_j = \sqrt{\frac{\sum_{n=l}^{N-j} (x_{n+j} - \bar{x}_{n,j})^2}{N - j - l + 1}}. \quad (7)$$

Here the index  $j$  denotes the forecast lead time in months,  $x_{n+j}$  is a true value of the predicted variable at time  $t_{n+j}$ ,  $\bar{x}_{n,j}$  is the value predicted by the model starting at time  $t_n$ .

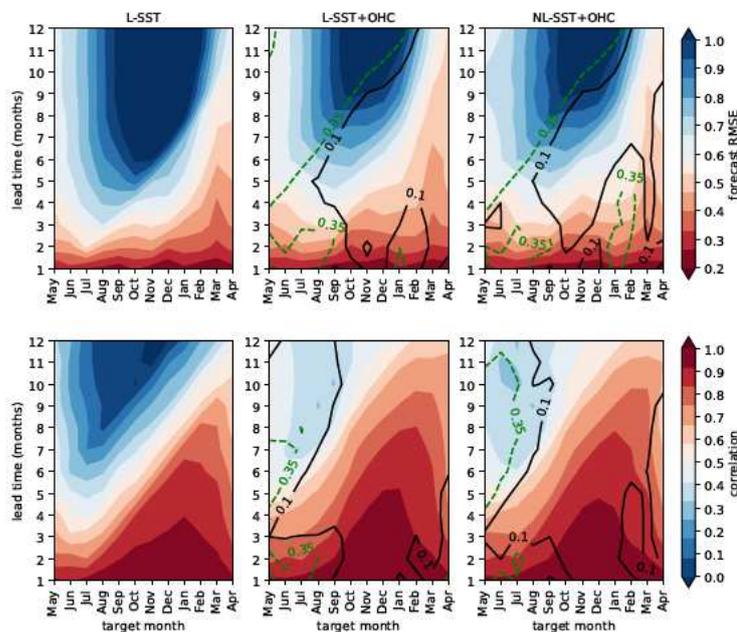
The second metric is the Pearson correlation between the variable and its forecast:

$$r_j = \frac{\sum_{n=l}^{N-j} \Delta x_{n+j} \cdot \Delta \bar{x}_{n,j}}{\sqrt{\sum_{n=l}^{N-j} (\Delta x_{n+j})^2 \cdot \sum_{n=l}^{N-j} (\Delta \bar{x}_{n,j})^2}}, \quad (8)$$

where  $\Delta x_{n+j}$  and  $\Delta \bar{x}_{n,j}$  are the deviations of  $x_{n+j}$  and  $\bar{x}_{n,j}$  from their means. The metrics (7)-(8) complement each other: while the RMSE measures a distance between the real and predicted values, the correlation metric reflects their relative similarity (in terms of linear relationships).

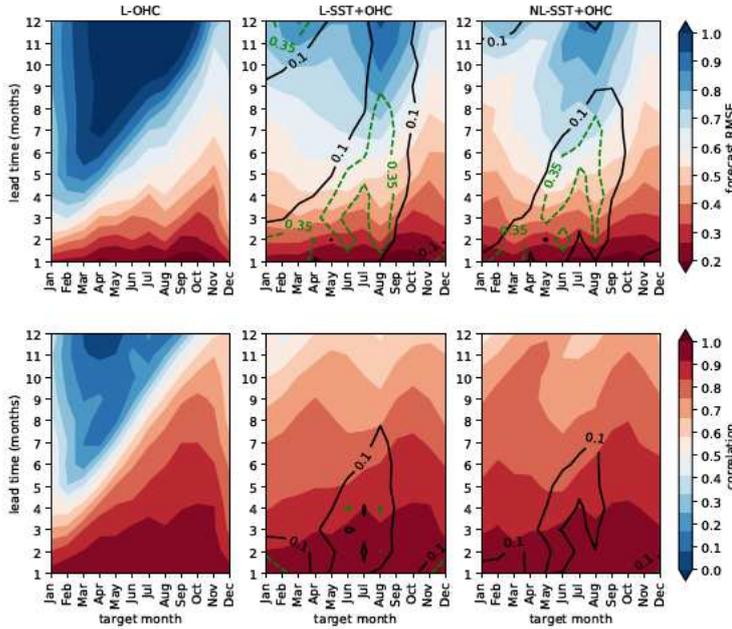
*Hindcast skill of SST field* In practice, the correct prediction of SST variability in the tropical Pacific is the main goal of both statistical and dynamical ENSO models (Barnston et al., 2012). Fig. 4 shows the spatial distributions of RMSE for all components  $\mathbf{x}_n$  of the SST field obtained using the three considered models described above. As we can see from Fig. 4, all models provide the best forecasts in the central tropical Pacific, for lead times up to 5 months. At the same time, both the L-SST+OHC and NL-SST+OHC models yield significantly lower the RMSE than the L-SST model, for lead times up to 11 months.

To find the areas where these SST+OHC models demonstrate statistically significant improvements of the prediction skills, we use a surrogates test similar to the test suggested by Mukhin et al. (2021). First, we produce 1000 surrogates of the first SSTA PC using the optimal L-SST model and 100 surrogates



**Fig. 5** Seasonal dependence of the prediction skills of different models calculated for the Niño 3.4 index. The forecast RMSE 7 normalized by the standard deviation of the  $CO_2$  detrended Niño 3.4 index (upper panels) and the correlations 8 (bottom panels) are shown for different target months and lead times ranging from 1 to 12 months. The contours in the central and right panels bound the areas of significant improvements of the joint L-SST+OHC and NL-SST+OHC model prediction skills relative to the separate L-SST model. The left-tailed statistical test is used for the RMSE metric, and the right-tailed test is used for the correlation metric (see the text).

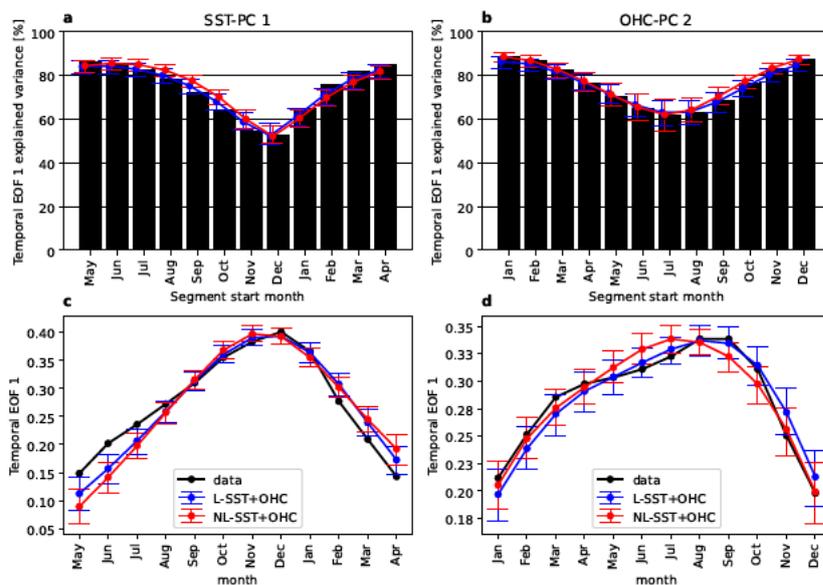
297 of the second OHCA PC using the optimal L-OHC model. Each surrogate is  
 298 a stochastic time series produced by the corresponding model starting from  
 299 random initial point. The length of each surrogate  $N = 732$  months is equal to  
 300 the length of the original dataset. Then we train the SST+OHC models with  
 301 optimal values of their hyperparameters on each pair of surrogates and calcu-  
 302 late the metric (7) in the physical space. Using the obtained ensemble of the  
 303 RMSE values we can find the areas where the RMSE of the SST+OHC models  
 304 constructed from data lies on the tail of distribution. Thus the null hypothesis  
 305 to be rejected supposes that the model that includes the information about  
 306 both SST and OHC variability delivers the same prediction skills as the sepa-  
 307 rate SST and OHC models. The areas where the null hypothesis is rejected at  
 308 significance levels of 0.1 and 0.35 are marked by contours in Fig. 4. We observe  
 309 that the the most significant improvements of the 11-month prediction skills  
 310 using the SST+OHC models appears in the central tropical Pacific around the  
 311 Niño 3.4 region (5S-5N, 160E-150W) (Bamston et al., 1997).



**Fig. 6** The same as in Fig. 5 but for the OHCA PC2 (see the text). The RMSE is normalized to the standard deviation of the OHCA PC2.

312 *Seasonal dependence of model prediction skills* Fig. 5 shows the month-to-  
 313 month distribution of the prediction skill of different models for the Niño 3.4  
 314 index. The top panel corresponds to the RMSE metric (7) and the low panel  
 315 – to the correlation metric (8). For all models we observe drop of the skills  
 316 in the Niño 3.4 forecasts in the late spring and summer months. In other  
 317 words, the Niño 3.4 index is less-predictable in the months when ENSO events  
 318 normally start to develop. It is a manifestation of the so-called "spring pre-  
 319 dictability barrier" which is a common problem for statistical and dynamical  
 320 ENSO models (Bamston et al., 1997). From Fig. 5 one can see that the joint  
 321 SST+OHC models have significantly better prediction skills as compared with  
 322 the separate linear SST model, for all months including ones associated with  
 323 the spring barrier. The 0.1 and 0.35 significance levels are evaluated for both  
 324 metrics using the statistical test described above with 1000 surrogates. Fig.  
 325 6 displays the month-to-month dependence of the prediction skills for second  
 326 OHCA PC, which plays the role of an index characterizing OHC accumulation  
 327 in the tropical Pacific. For this variable, we also see the predictability barrier,  
 328 but now the drop of the prediction skills fall in the winter months. Again, the  
 329 SST+OHC models outperform the separate OHC-based model.

330 Thus we found that both the L-SST+OHC and NL-SST+OHC models  
 331 yield significant benefits in the forecast as compared with separate models.



**Fig. 7** Leading temporal EOF in ENSO dynamics. Top panel: variance explained by the leading temporal (12-month) EOF of the SSTA PC1 (a) and the OHCA PC2 (b) is shown as a function of the first month of the EOF. The black boxes correspond to the data-based PCs and the color curves correspond to the PCs produced by the L-SST+OHC (blue) and NL-SST+OHC (red) models. Bottom panel: the leading temporal EOF of the SSTA PC1 (c) and the OHCA PC2 (d). The black curves correspond to the data-based temporal EOFs and the color curves correspond to the temporal EOFs in the behavior of the L-SST+OHC (blue) and NL-SST+OHC (red) models. The 90% confidence intervals are evaluated using the ensemble of 1000 model runs (see the text).

332 At the same time, they deliver almost the same prediction skills, with no  
 333 significant differences. This means that the nonlinearity does not matter for  
 334 intra-annual multi-month forecasts in the tropical Pacific region. In the next  
 335 section, we show that, nevertheless, the NL-SST+OHC model captures sig-  
 336 nificant nonlinear laws that manifest themselves in the observed dynamics on  
 337 inter-annual scales.

### 338 3.3.2 Simulation of ENSO phase locking

339 The spring predictability barrier is closely connected with ENSO phase locking  
 340 to the annual cycle (Liu et al., 2019). Tippett and L'Heureux (2020) demon-  
 341 strated that approximately 90% of observed seasonal evolution of the Niño  
 342 3.4 index can be explained by deterministic seasonal signal defined on the  
 343 June–May interval, which reaches a maximum in December and has the low-

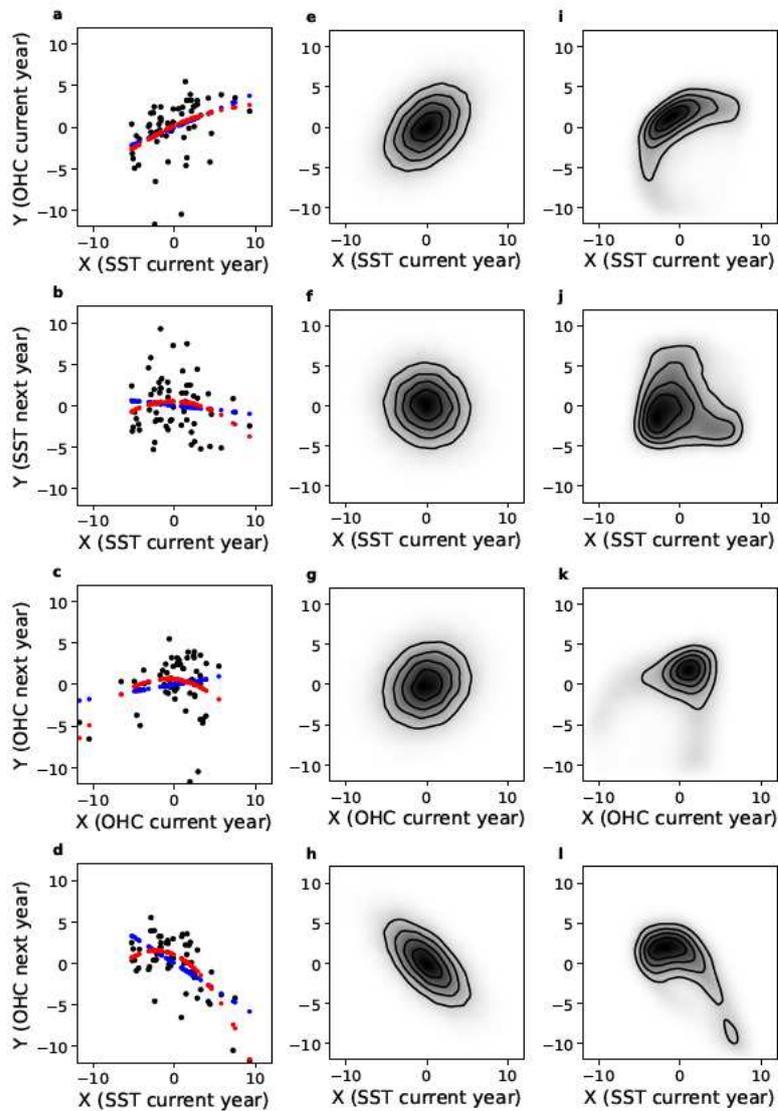
est absolute values at the boundaries of the interval – in June and May of the following year. This signal, multiplied by different amplitudes in different Jun-May windows, “isolates the intrinsic seasonal cycle of ENSO evolution and its phase-locking to the annual cycle” (Tippett and L’Heureux, 2020).

Technically, we can retrieve the seasonal cycle of this type from a monthly ENSO index by means of the EOF decomposition applied to the set of non-overlapping successive 12-month segments of the index time series. The leading EOF of the obtained yearly 12-channel time series (hereinafter, the temporal EOF) determines the required 12-month cycle, whereas the corresponding PC is a yearly time series of the cycle amplitudes. Obviously, this leading EOF depends on the dividing the time series into the segments, or, equivalently, selecting the start month of the segment. Since we naturally interest in obtaining the cycle that explains a substantial part of variability, we select the start month providing that the leading temporal EOF captures the largest variance of the original index.

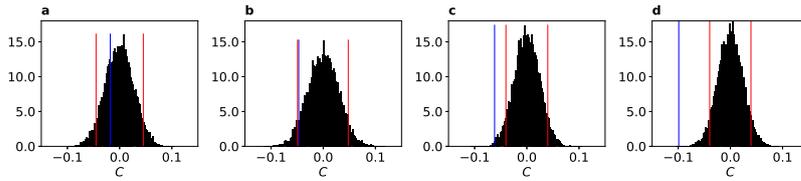
We have checked if strong seasonal cycles underlie the first SSTA PC and second OHCA PC time series. The black boxes in Fig. 7 indicate the fraction of variance explained by the leading temporal EOFs depending on the segment start months. This figure shows a strong cycle in the SSTA PC that starts in May-June and captures about 86% of variance, which is in agreement with the results obtained by Tippett and L’Heureux (2020) for the Niño 3.4 index. For the second OHCA PC, 88% of variance is explained by the cycle starting from December-January. This tells us that we would be facing a winter (not spring) barrier, if we constructed a model based on this OHC time series alone. In Fig. 7 (c)-(d) the temporal EOFs determining the shapes of the above cycles are plotted. Although, as expected, the SST cycle peaks in December, the OHC accumulation cycle culminates in August-September.

Now let us look how our data-driven models reproduce these cycles. We repeated the cycle analysis described above for the time series generated by both the L-SST+OHC and NL-SST+OHC models; the results are shown in Fig. 7 by blue and red, respectively. We performed 1000 model runs per model, calculated the leading temporal EOFs for each time series from this ensemble, and then evaluated the confidence intervals for the EOFs and variances in each month. Overall, we can say that both models reproduce well the temporal EOF patterns and therefore capture the seasonal cycles in the two key variables of ENSO.

Fig. 8 (a)-(d) shows several planes of lead-lag and synchronous dependencies between the temporal EOF (cycle) amplitudes in the SSTA PC1 and the OHCA PC2 time series. It can be observed from the planes (c) and (d) that the dependencies of the OHC cycle amplitude on the previous OHC and SST cycle amplitudes look nonlinear. To verify the nonlinearities observed, we fit the linear  $Y = B \cdot X + A + \varepsilon$  as well as quadratic  $Y = C \cdot X^2 + B \cdot X + A + \varepsilon$  functions to the observed dependencies and analyze the significance of the quadratic terms. The traditional least square method was used for estimating the coefficients  $A$ ,  $B$  and  $C$  as well as the variance of an approximation error  $\varepsilon$  represented as Gaussian noise without point-to-point correlations. The re-



**Fig. 8** Relationships between state variables. Left column: the planes of the amplitudes of the leading temporal EOF of the SSTA PC1 (May is the first month) and the OHCA PC2 (January is the first month). Black points correspond to data and colored points – to the approximations of data via linear (blue) and quadratic regression models (red). Central and right columns: PDs in the same planes, estimated from 1000 runs of the L-SST+OHC and NL-SST+OHC models, respectively.



**Fig. 9** Testing the significance of the quadratic term. PDs of the coefficient  $C$  of the quadratic model fitted to each of 1000 surrogate time series produced by the linear model. Each plot from left to right corresponds to a particular plane from Fig. 8 (a)-(d). Blue lines denote the values of coefficient  $C$  of the quadratic model fitted on data. Red lines mark the 10th and 90th percentiles.

390 sulting fits are shown by blue and red in Fig. 8 (a)-(d). We can notice that the  
 391 curves of the linear and quadratic models are most distinct in the planes (c)-  
 392 (d). Testing the significance of the quadratic approximation can be performed  
 393 via rejecting the null-hypothesis that the obtained value of the coefficient  $C$   
 394 in the quadratic term could be obtained from a similar sample but with a linear  
 395 dependence between variables. For each plane, using the linear function fitted  
 396 to the original sample, we generated an ensemble of 1000 random surrogate  
 397 samples. Then we fitted the quadratic model to each surrogate and used the  
 398 resulting values of  $C$  as the ensemble corresponding to the null-hypothesis.  
 399 Such ensembles relating to the planes from Fig. 8 are shown in Fig. 9. It is  
 400 seen from this figure that the quadratic approximation is significant by level  
 401 0.1 for the planes (c) and (d) from Fig. 8 indicating an apparent nonlinear  
 402 dependence of the current OHC cycle on the previous OHC and SST cycles.

403 Next, we can use our optimal data-driven models for verifying the detected  
 404 nonlinear relationships. To this end, we took an ensemble of 1000 monthly time  
 405 series of SSTA PC1 and OHCA PC2 generated by the NL-SST+OHC model,  
 406 and, for comparison, the same ensemble but generated by the L-SST+OHC  
 407 model. Then we calculated the OHC and SST cycles amplitudes from these  
 408 time series and plotted resulting probability densities (PDs) in the planes  
 409 shown in Fig. 8. Naturally, no nonlinearity can be captured by a linear model,  
 410 therefore, the L-SST+OHC model yields Gaussian PD in all the planes consid-  
 411 ered. However, the optimal nonlinear (NL-SST+OHC) model produces appar-  
 412 ently non-Gaussian PDs thus confirming pronounce nonlinear laws underlying  
 413 the inter-cycle dynamics. We can also conclude that although this nonlinear  
 414 model does not provide additional benefits in short-term forecasting over the  
 415 linear model, it nevertheless more adequately reflects the dynamical properties  
 416 of ENSO on interannual scales.

## 417 4 Discussion

418 In this study we have utilized gridded datasets of the tropical Pacific SST and  
 419 OHC anomalies in the 0-300 m depth layer to reveal the dynamical variables  
 420 containing meaningful information about ENSO as well as to construct the

421 data-driven model based on these variables. The EOF analysis applied to both  
422 data sets clearly demonstrates phase relationships between the first SSTA and  
423 second OHCA PCs yielding the largest absolute cross-correlations when the  
424 corresponding time series are shifted by about 5-9 months. While the first  
425 SSTA EOF is known to be associated with SST variability in the highly ENSO-  
426 related region, the second OHCA EOF, in accordance with Clarke et al. (2007),  
427 likely reflects the OHC accumulation and discharge before and during the El  
428 Niño events, respectively.

429 We constructed and compared different (linear and nonlinear) data-driven  
430 stochastic models based on the SSTA PC1 and OHCA PC2 variables taken  
431 separately and together. It is shown that the data-driven models combining  
432 these two variables yield significant benefits in predicting both the SST and  
433 OHC variability and allow to substantially lower the seasonal predictability  
434 barriers as compared with the separate models. Thus the second OHCA EOF  
435 can be used as an effective additional ENSO predictor in statistical models.

436 We then obtained that the seasonal cycles (dominating 12-month patterns)  
437 in SST and OHC variability are different: while the SST cycle peaks in early  
438 winter and drops in late spring, the OHC accumulation cycle is shifted forward  
439 by approximately 8 months. Generally speaking, a strong seasonal cycle in a  
440 single variable, defined as the leading temporal EOF, unavoidably leads to the  
441 existence of a predictability barrier when we use a statistical model derived  
442 from the time series of this variable. The reason is that the variable values  
443 at months inside the cycle interval are highly correlated, but the inter-cycle  
444 connections are more stochastic. A possible way to overcome such a barrier is to  
445 invoke an additional variable that is connected with the original one, but has no  
446 barrier in the same months. As it is seen from figures 5–7, there is a pronounce  
447 winter (Dec-Jan) predictability barrier in the OHC accumulation variability, in  
448 contrast to the well-known spring barrier in the SST variability. Note that if the  
449 similar seasonal patterns in the SST-based Niño-family indices are mentioned  
450 in other studies (Kondrashov et al., 2005; Tippett and L’Heureux, 2020; Chen  
451 and Jin, 2020, 2021), the corresponding OHC seasonal evolution has not been  
452 in focus yet. Since it is found that the joint SST+OHC models outperform  
453 the separate SST and OHC models in prediction skill, we conclude that the  
454 detected SST and OHC accumulation cycles strongly interact, and hence, the  
455 use of the combined SST-OHC phase space helps to lower the seasonal barriers  
456 in the ENSO variables.

457 We also derived from data that the inter-annual interaction of the cycles is  
458 substantially nonlinear, and the optimal data-driven model with the nonlinear  
459 parameterization confirms this. It is important that ENSO manifests its non-  
460 linear dynamical properties on long, interannual scales, while nonlinearity on  
461 several-month intervals is not resolved. This finding opens prospects for de-  
462 veloping nonlinear statistical models for yearly ENSO variability, which could  
463 expand the horizon of ENSO forecasts.

464 **Acknowledgements** This research was supported by grant #19-42-04121 from the Rus-  
465 sian Science Foundation (selecting the state variables and constructing the data-driven mod-

466 els of ENSO). Also, it was supported by grant #19-02-00502 from the Russian Foundation  
 467 for Basic Research (studying the ENSO phase locking to the annual cycle). The authors are  
 468 grateful to professor Alexander M. Feigin as well as Dr. Andrey Gavrilov and Dr. Eugeny  
 469 Loskutov from the Institute of Applied Physics of RAS, for fruitful discussions concerning  
 470 this research.

## 471 A Appendix: Bayesian approach to ENSO model learning and 472 optimization

473 Here we outline the Bayesian approach we use for learning and optimization of the stochastic  
 474 model (2). The optimal model relied on observed data is supposed to be a right balance  
 475 between the “too simple” model poorly describing data and “too complex” model which  
 476 contains too many parameters and tends to be overfitted to the available sample rather than  
 477 to capture the laws underlying the dynamics. Let the  $\mathbf{H} = \{H_1, H_2, \dots, H_i, \dots\}$  is the set  
 478 of possible hypotheses about the model complexity. In the case of a stochastic model (2)  
 479 each hypothesis  $H_i$  is determined by the particular combination of the hyperparameters  $l, q$   
 480 in the case of the linear parameterization (3)-(4) of deterministic part  $\mathbf{f}$  and  $l, m$  in the case  
 481 of the nonlinear parameterization (5). According to the Bayes rule the probability  $P(H_i|\mathbf{Y})$   
 482 that the model  $H_i$  produces the observed time series  $\mathbf{Y} = (\mathbf{p}_1, \dots, \mathbf{p}_n)$  is equal to:

$$P(H_i|\mathbf{Y}) = \frac{P(\mathbf{Y}|H_i)P(H_i)}{\sum_i P(\mathbf{Y}|H_i)P(H_i)}. \quad (\text{A1})$$

483 Here probability density function (PDF)  $P(\mathbf{Y}|H_i)$  is the evidence (marginal likelihood) of  
 484 the model  $H_i$  characterizing the probability of the observed data  $\mathbf{Y}$  to belong to the whole  
 485 possible ensemble of time series which can be produced by the model  $H_i$ ;  $P(H_i)$  is a prior  
 486 probability of the model  $H_i$ . The denominator in (A1) is a normalization term which does  
 487 not depend on  $H_i$ . Assuming all the models from  $\mathbf{H}$  equiprobable a priori, the expression  
 488 (A1) can be rewritten as  $P(H_i|\mathbf{Y}) = \alpha P(\mathbf{Y}|H_i)$  where  $\alpha$  is independent of  $H_i$ . Let us define  
 489 the Bayesian criterion of the model optimality

$$L = -\log P(\mathbf{Y}|H_i), \quad (\text{A2})$$

490 minimization of which leads to maximization of the PDF  $P(\mathbf{Y}|H_i)$ . The optimality criterion  
 491 (A2) has a clear interpretation. If the model  $H_i$  is too simple, than the observed data likely  
 492 lie on a tail of the PDF  $P(\mathbf{Y}|H_i)$ . Therefore, the probability that the observed data  $\mathbf{Y}$  could  
 493 be produced by such a model is small. In contrast, the overfitted model, due to a large  
 494 number of parameters, produces a widely distributed population of different datasets, which  
 495 lowers again the PDF of the observed  $\mathbf{Y}$ . Therefore, the optimality (A2) helps to select the  
 496 optimal model that is neither too simple nor overfitted model.

497 The evidence  $P(\mathbf{Y}|H_i)$  is expressed via integration of the product of the corresponding  
 498 likelihood function  $P(\mathbf{Y}|\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}, H_i)$  and the prior distribution  $P(\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}|H_i)$  over the model  
 499 parameter space:

$$P(\mathbf{Y}|H_i) = \int P(\mathbf{Y}|\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}, H_i) \cdot P(\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}|H_i) d\mu_{\mathbf{f}} \mu_{\widehat{\mathbf{g}}}. \quad (\text{A3})$$

500 Here the vectors  $\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}$  contain parameters of the deterministic part  $\mathbf{f}$  and the stochastic  
 501 part of the model (2), respectively. The likelihood function  $P(\mathbf{Y}|\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}, H_i)$  corresponds to  
 502 the assumption that the stochastic part of the model is the delta-correlated in time Gaussian  
 503 process with the amplitude  $\widehat{\mathbf{g}}$  (see Sec. 2.2.1):

$$P(\mathbf{Y}|\mu_{\mathbf{f}}, \mu_{\widehat{\mathbf{g}}}, H_i) = \prod_{n=1}^l P_{\mathcal{N}}(\mathbf{p}_n, \widehat{\mathbf{I}}) \times \prod_{n=l+1}^N P_{\mathcal{N}}(\mathbf{p}_n - \mathbf{f}(\mathbf{p}_{n-1}, \dots, \mathbf{p}_{n-l}), \widehat{\mathbf{g}}\widehat{\mathbf{g}}^T). \quad (\text{A4})$$

504 Here  $P_{\mathcal{N}}(\mathbf{u}, \widehat{\Sigma}) := \frac{1}{\sqrt{(2\pi)^d |\widehat{\Sigma}|}} \exp\left(-\frac{1}{2} \mathbf{u}^T \widehat{\Sigma}^{-1} \mathbf{u}\right)$ ,  $\mathbf{u} \in \mathbb{R}^d$ ,  $\prod_{n=1}^l P_{\mathcal{N}}(\mathbf{p}_n, \widehat{\mathbf{I}})$  is a term describ-  
 505 ing PDF of the initial state of the model (see (Gavrilov et al., 2017) for more details). The

506 prior PDF  $P(\boldsymbol{\mu}_f, \boldsymbol{\mu}_g | H_i)$  is the product of Gaussian PDFs for each parameter of the model.  
 507 The proper choice of dispersions of the corresponding PDFs for linear (3)-(4) and nonlinear  
 508 (5) parametrizations is discussed in detail in Mukhin et al. (2021); Seleznev et al. (2019).

509 The evidence (A3) is estimated using the Laplace's method based on approximate inte-  
 510 grating in the neighborhood of maximum of integrand. Let us denote the minus logarithm  
 511 of the integrand in (A3) as  $\Psi_{H_i}(\boldsymbol{\mu}_f, \boldsymbol{\mu}_g)$ . Then the integrand can be rewritten as:

$$P(\mathbf{Y} | \boldsymbol{\mu}_f, \boldsymbol{\mu}_g, H_i) \cdot P(\boldsymbol{\mu}_f, \boldsymbol{\mu}_g | H_i) = \exp(-\Psi_{H_i}(\boldsymbol{\mu}_f, \boldsymbol{\mu}_g)). \quad (\text{A5})$$

512 The integration of (A3) using Laplace method by decomposing the function  $\Psi_{H_i}(\boldsymbol{\mu}_f, \boldsymbol{\mu}_g)$  in  
 513 the neighborhood of its minimum into a second-order Taylor series leads to the following  
 514 expression for the optimality criterion (A2):

$$L = -\log P(\mathbf{Y} | H_i) = \Psi_{H_i}(\overline{\boldsymbol{\mu}_f}, \overline{\boldsymbol{\mu}_g}) + \frac{1}{2} \ln \left[ \frac{1}{2\pi} (\nabla \nabla^T \Psi_{H_i}(\overline{\boldsymbol{\mu}_f}, \overline{\boldsymbol{\mu}_g})) \right]. \quad (\text{A6})$$

515 Here  $\Psi_{H_i}(\overline{\boldsymbol{\mu}_f}, \overline{\boldsymbol{\mu}_g})$  is the function value at its minimum;  $\nabla \nabla^T \Psi_{H_i}(\overline{\boldsymbol{\mu}_f}, \overline{\boldsymbol{\mu}_g})$  is the matrix of  
 516 the second derivatives (hessian matrix) at the minimum. The first term in (A6) reflects the  
 517 accuracy of data approximation by the model. It decreases with expanding the model com-  
 518 plexity, i.e. with growing of number of parameters, and therefore prevents too simple models.  
 519 In contrast, the second term in (A6) increases with growing of the number of model param-  
 520 eters and penalizes the overfitted models. The particular algorithm we use for numerical  
 521 calculation of (A6) can be found in (Seleznev et al., 2019).

## 522 References

- 523 Alexander MA, Bladé I, Newman M, Lanzante JR, Lau NC, Scott JD (2002) The  
 524 Atmospheric Bridge: The Influence of ENSO Teleconnections on Air–Sea Interac-  
 525 tion over the Global Oceans. *Journal of Climate* 15(16):2205–2231, doi:10.1175/1520-  
 526 0442(2002)015<2205:TABTIO>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
 527 journals/clim/15/16/1520-0442\\_2002\\_015\\_2205\\_tabtio\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/clim/15/16/1520-0442_2002_015_2205_tabtio_2.0.co_2.xml)
- 528 Anderson DLT, McCreary JP (1985) Slowly Propagating Disturbances in a Coupled Ocean-  
 529 Atmosphere Model. *Journal of Atmospheric Sciences* 42(6):615–630, doi:10.1175/1520-  
 530 0469(1985)042<0615:SPDIAC>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
 531 journals/atsc/42/6/1520-0469\\_1985\\_042\\_0615\\_spdiac\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/atsc/42/6/1520-0469_1985_042_0615_spdiac_2_0_co_2.xml)
- 532 Bamston AG, Chelliah M, Goldenberg SB (1997) Documentation of a highly ENSO-  
 533 related sst region in the equatorial pacific: Research note. *Atmosphere-Ocean* 35(3):367–  
 534 383, doi:10.1080/07055900.1997.9649597, URL [https://doi.org/10.1080/07055900.  
 535 1997.9649597](https://doi.org/10.1080/07055900.1997.9649597)
- 536 Barnston AG, Tippett MK, L'Heureux ML, Li S, Dewitt DG (2012) Skill of real-time sea-  
 537 sonal ENSO model predictions during 2002-11: Is our capability increasing? *Bulletin of  
 538 the American Meteorological Society* 93(5):631–651, doi:10.1175/BAMS-D-11-00111.1
- 539 Bjerknes J (1969) Monthly Weather Review Atmospheric Teleconnections From the Equa-  
 540 torial Pacific. *Monthly Weather Review* 97(3):163–172, URL [http://journals.ametsoc.  
 541 org/doi/abs/10.1175/1520-0493\(1969\)0973C0163:ATFTEP3E2.3.CO;2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0493(1969)0973C0163:ATFTEP3E2.3.CO;2)
- 542 Byshev VI, Neiman VG, Romanov YA, Serykh IV, Sonechkin DM (2016) Statistical signifi-  
 543 cance and climatic role of the Global Atmospheric Oscillation. *Oceanology* 56(2):165–171,  
 544 doi:10.1134/S000143701602003X
- 545 Cane MA, Zebiak SE (1985) A Theory for El Niño and the Southern Oscillation.  
 546 *Science* 228(4703):1085–1087, doi:10.1126/science.228.4703.1085
- 547 Chen HC, Jin FF (2020) Fundamental Behavior of ENSO Phase Locking. *Journal of Climate*  
 548 33(5):1953–1968, doi:10.1175/JCLI-D-19-0264.1, URL [https://journals.ametsoc.org/  
 549 view/journals/clim/33/5/jcli-d-19-0264.1.xml](https://journals.ametsoc.org/view/journals/clim/33/5/jcli-d-19-0264.1.xml)
- 550 Chen HC, Jin FF (2021) Simulations of ENSO Phase-Locking in CMIP5 and CMIP6. *Journal  
 551 of Climate* 34(12):5135–5149, doi:10.1175/JCLI-D-20-0874.1, URL [https://journals.  
 552 ametsoc.org/view/journals/clim/34/12/JCLI-D-20-0874.1.xml](https://journals.ametsoc.org/view/journals/clim/34/12/JCLI-D-20-0874.1.xml)

- 553 Chen HC, Tseng YH, Hu ZZ, Ding R (2020) Enhancing the ENSO Predictability beyond  
554 the Spring Barrier. *Scientific Reports* 10(1):1–12, doi:10.1038/s41598-020-57853-7
- 555 Chen S, Wu R, Chen W, Yu B, Cao X (2016) Genesis of westerly wind bursts over the  
556 equatorial western Pacific during the onset of the strong 2015–2016 El Niño. *Atmospheric*  
557 *Science Letters* 17(7):384–391, doi:https://doi.org/10.1002/asl.669, URL <https://doi.org/10.1002/asl.669>
- 558 Cheng L, Trenberth KE, Fasullo J, Boyer T, Abraham J, Zhu J (2017) Improved estimates of  
559 ocean heat content from 1960 to 2015. *Science Advances* 3(3), doi:10.1126/sciadv.1601545,  
560 URL <https://advances.sciencemag.org/content/3/3/e1601545>
- 561 Cheng L, Trenberth KE, Fasullo JT, Mayer M, Balmaseda M, Zhu J (2019) Evolution  
562 of Ocean Heat Content Related to ENSO. *Journal of Climate* 32(12):3529–3556,  
563 doi:10.1175/JCLI-D-18-0607.1, URL <https://journals.ametsoc.org/view/journals/clim/32/12/jcli-d-18-0607.1.xml>
- 564 Chiodi AM, Harrison DE, Vecchi GA (2014) Subseasonal Atmospheric Variability and El  
565 Niño Waveguide Warming: Observed Effects of the Madden–Julian Oscillation and Westerly  
566 Wind Events. *Journal of Climate* 27(10):3619–3642, doi:10.1175/JCLI-D-13-00547.1,  
567 URL <https://journals.ametsoc.org/view/journals/clim/27/10/jcli-d-13-00547.1.xml>
- 568 Clarke AJ, Van Gorder S (2003) Improving El Niño prediction using a space-time inte-  
569 gration of Indo-Pacific winds and equatorial Pacific upper ocean heat content. *Geo-*  
570 *physical Research Letters* 30(7), doi:https://doi.org/10.1029/2002GL016673, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2002GL016673>
- 571 Clarke AJ, Gorder SV, Colantuono G (2007) Wind Stress Curl and ENSO Dis-  
572 charge/Recharge in the Equatorial Pacific. *Journal of Physical Oceanography* 37(4):1077–  
573 1091, doi:10.1175/JPO3035.1, URL <https://journals.ametsoc.org/view/journals/phoc/37/4/jpo3035.1.xml>
- 574 Cybenko G (1989) Approximations by superpositions of sigmoidal functions. *Approximation*  
575 *Theory and its Applications* 9(3):17–28, doi:10.1007/BF02836480
- 576 Deser C, Alexander MA, Xie SP, Phillips AS (2009) Sea Surface Temperature  
577 Variability: Patterns and Mechanisms. *Annual Review of Marine Science*  
578 2(1):115–143, doi:10.1146/annurev-marine-120408-151453, URL <https://doi.org/10.1146/annurev-marine-120408-151453>
- 579 Fedorov AV, Harper SL, Philander SG, Winter B, Wittenberg A (2003) How Pre-  
580 dictable is El Niño? *Bulletin of the American Meteorological Society* 84(7):911–920,  
581 doi:10.1175/BAMS-84-7-911, URL <https://journals.ametsoc.org/view/journals/bams/84/7/bams-84-7-911.xml>
- 582 Galanti E, Tziperman E (2000) ENSO’s Phase Locking to the Seasonal Cycle in the  
583 Fast-SST, Fast-Wave, and Mixed-Mode Regimes. *Journal of the Atmospheric Sci-*  
584 *ences* 57(17):2936–2950, doi:10.1175/1520-0469(2000)057<2936:ESPLTT>2.0.CO;2,  
585 URL [https://doi.org/10.1175/1520-0469\(2000\)057<2936:ESPLTT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<2936:ESPLTT>2.0.CO;2), [https://journals.ametsoc.org/jas/article-pdf/57/17/2936/3449342/1520-0469\(2000\)057\\_2936\\_espltt\\_2\\_0\\_co\\_2.pdf](https://journals.ametsoc.org/jas/article-pdf/57/17/2936/3449342/1520-0469(2000)057_2936_espltt_2_0_co_2.pdf)
- 586 Gavrilov A, Loskutov E, Mukhin D (2017) Bayesian optimization of empirical model  
587 with state-dependent stochastic forcing. *Chaos, Solitons and Fractals* 104:327–337,  
588 doi:10.1016/j.chaos.2017.08.032, URL <http://dx.doi.org/10.1016/j.chaos.2017.08.032>
- 589 Gavrilov A, Seleznev A, Mukhin D, Loskutov E, Feigin A, Kurths J (2019) Linear dynam-  
590 ical modes as new variables for data-driven ENSO forecast. *Climate Dynamics* 52(3-4),  
591 doi:10.1007/s00382-018-4255-7
- 592 Hannachi A, Jolliffe IT, Stephenson DB (2007) Empirical orthogonal functions and re-  
593 lated techniques in atmospheric science: A review. *International Journal of Climatology*  
594 27(9):1119–1152, doi:10.1002/joc.1499, URL <http://doi.wiley.com/10.1002/joc.1499>
- 595 Hu S, Fedorov AV (2019) The extreme El Niño of 2015–2016: the role of westerly and  
596 easterly wind bursts, and preconditioning by the failed 2014 event. *Climate Dynamics*  
597 52(12):7339–7357, doi:10.1007/s00382-017-3531-2
- 598 Huang B, Thorne PW, Banzon VF, Boyer T, Chepurin G, Lawrimore JH, Menne  
599 MJ, Smith TM, Vose RS, Zhang HM (2017) Extended Reconstructed Sea  
600 Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Inter-

- 611 comparisons. *Journal of Climate* 30(20):8179–8205, doi:10.1175/JCLI-D-16-0836.1,  
612 URL <https://doi.org/10.1175/JCLI-D-16-0836.1>, [https://journals.ametsoc.org/  
613 jcli/article-pdf/30/20/8179/4680731/jcli-d-16-0836\\_1.pdf](https://journals.ametsoc.org/jcli/article-pdf/30/20/8179/4680731/jcli-d-16-0836_1.pdf)
- 614 Jin EK, Kinter JL, Wang B, Park CK, Kang IS, Kirtman BP, Kug JS, Kumar A, Luo  
615 JJ, Schemm J, Shukla J, Yamagata T (2008) Current status of ENSO prediction skill in  
616 coupled ocean-atmosphere models. *Climate Dynamics* 31(6):647–664, doi:10.1007/s00382-  
617 008-0397-3
- 618 Jin FF (1997) An Equatorial Ocean Recharge Paradigm for ENSO. Part I: Concep-  
619 tual Model. *Journal of the Atmospheric Sciences* 54(7):811–829, doi:10.1175/1520-  
620 0469(1997)054<0811:AEORPF>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
621 journals/atsc/54/7/1520-0469\\_1997\\_054\\_0811\\_aeorpf\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/atsc/54/7/1520-0469_1997_054_0811_aeorpf_2.0.co_2.xml)
- 622 Jin FF, Neelin JD (1993) Modes of Interannual Tropical Ocean–Atmosphere Interac-  
623 tion—a Unified View. Part I: Numerical Results. *Journal of Atmospheric Sciences*  
624 50(21):3477–3503, doi:10.1175/1520-0469(1993)050<3477:MOITOI>2.0.CO;2, URL  
625 [https://journals.ametsoc.org/view/journals/atsc/50/21/1520-0469\\_1993\\_050\\_  
626 3477\\_moittoi\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/atsc/50/21/1520-0469_1993_050_3477_moittoi_2.0.co_2.xml)
- 627 Kondrashov D, Kravtsov S, Robertson AW, Ghil M (2005) A hierarchy of data-based  
628 ENSO models. *Journal of Climate* 18(21):4425–4444, doi:10.1175/JCLI3567.1, URL [http:  
629 //journals.ametsoc.org/doi/pdf/10.1175/JCLI3567.1](http://journals.ametsoc.org/doi/pdf/10.1175/JCLI3567.1)
- 630 Levine AFZ, Jin FF (2017) A simple approach to quantifying the noise–ENSO interaction.  
631 Part I: deducing the state-dependency of the windstress forcing using monthly mean  
632 data. *Climate Dynamics* 48(1):1–18, doi:10.1007/s00382-015-2748-1, URL [https://doi.  
633 org/10.1007/s00382-015-2748-1](https://doi.org/10.1007/s00382-015-2748-1)
- 634 Liu Z, Jin Y, Rong X (2019) A Theory for the Seasonal Predictability Barrier: Threshold,  
635 Timing, and Intensity. *Journal of Climate* 32(2):423–443, doi:10.1175/JCLI-D-18-0383.1,  
636 URL [https://journals.ametsoc.org/view/journals/clim/32/2/jcli-d-18-0383.1.  
637 xml](https://journals.ametsoc.org/view/journals/clim/32/2/jcli-d-18-0383.1.xml)
- 638 Martinez-Villalobos C, Newman M, Vimont DJ, Penland C, David Neelin J (2019) Ob-  
639 served El Niño–La Niña Asymmetry in a Linear Model. *Geophysical Research Letters*  
640 46(16):9909–9919, doi:10.1029/2019GL082922
- 641 Meinen CS, McPhaden MJ (2000) Observations of Warm Water Volume Changes in the  
642 Equatorial Pacific and Their Relationship to El Niño and La Niña. *Journal of Climate*  
643 13(20):3551–3559, doi:10.1175/1520-0442(2000)013<3551:OOWWVC>2.0.CO;2, URL  
644 [https://journals.ametsoc.org/view/journals/clim/13/20/1520-0442\\_2000\\_013\\_  
645 3551\\_oowwvc\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/clim/13/20/1520-0442_2000_013_3551_oowwvc_2.0.co_2.xml)
- 646 Molkov YI, Loskutov EM, Mukhin DN, Feigin AM (2012) Random dynamical models  
647 from time series. *Physical Review E* 85(3):036216, doi:10.1103/PhysRevE.85.036216, URL  
648 <https://link.aps.org/doi/10.1103/PhysRevE.85.036216>
- 649 Mukhin D, Kondrashov D, Loskutov E, Gavrilov A, Feigin A, Ghil M (2015a) Predicting  
650 critical transitions in ENSO models. Part II: Spatially dependent models. *Journal of*  
651 *Climate* 28(5):1962–1976, doi:10.1175/JCLI-D-14-00240.1
- 652 Mukhin D, Loskutov E, Mukhina A, Feigin A, Zaliapin I, Ghil M (2015b) Predicting criti-  
653 cal transitions in ENSO Models. Part I: Methodology and simple models with memory.  
654 *Journal of Climate* 28(5):1940–1961, doi:10.1175/JCLI-D-14-00239.1
- 655 Mukhin D, Gavrilov A, Seleznev A, Buyanova M (2021) An Atmospheric Signal Lowering  
656 the Spring Predictability Barrier in Statistical ENSO Forecasts. *Geophysical Research*  
657 *Letters* 48(6):1–10, doi:10.1029/2020GL091287
- 658 Penland C, Sardeshmukh PD (1995) The Optimal Growth of Tropical Sea Sur-  
659 face Temperature Anomalies. *Journal of Climate* 8(8):1999–2024, doi:10.1175/1520-  
660 0442(1995)008<1999:TOGOTS>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
661 journals/clim/8/8/1520-0442\\_1995\\_008\\_1999\\_togots\\_2.0.co\\_2.xml](https://journals.ametsoc.org/view/journals/clim/8/8/1520-0442_1995_008_1999_togots_2.0.co_2.xml)
- 662 Philander SG, Fedorov A (2003) Is El Niño Sporadic or Cyclic? *Annual Review of Earth*  
663 *and Planetary Sciences* 31(1):579–594, doi:10.1146/annurev.earth.31.100901.141255, URL  
664 <https://doi.org/10.1146/annurev.earth.31.100901.141255>
- 665 Puy M, Vialard J, Lengaigne M, Guilyardi E (2016) Modulation of equatorial Pacific west-  
666 erly/easterly wind events by the Madden–Julian oscillation and convectively-coupled  
667 Rossby waves. *Climate Dynamics* 46(7):2155–2178, doi:10.1007/s00382-015-2695-x, URL  
668 <https://doi.org/10.1007/s00382-015-2695-x>

- 669 Seleznev A, Mukhin D, Gavrilov A, Loskutov E, Feigin A (2019) Bayesian framework for  
 670 simulation of dynamical systems from multidimensional data using recurrent neural net-  
 671 work. *Chaos* 29(12), doi:10.1063/1.5128372
- 672 Suarez MJ, Schopf PS (1988) A Delayed Action Oscillator for ENSO.  
 673 *Journal of the Atmospheric Sciences* 45(21):3283–3287, doi:10.1175/1520-  
 674 0469(1988)045<3283:ADAOFE>2.0.CO;2, URL [https://doi.org/10.1175/  
 675 1520-0469\(1988\)045<3283:ADAOFE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1988)045<3283:ADAOFE>2.0.CO;2), [https://journals.ametsoc.org/jas/  
 676 article-pdf/45/21/3283/3424429/1520-0469\(1988\)045\\_3283\\_adaofe\\_2\\_0\\_co\\_2.pdf](https://journals.ametsoc.org/jas/article-pdf/45/21/3283/3424429/1520-0469(1988)045_3283_adaofe_2_0_co_2.pdf)
- 677 Sullivan A, Zhong W, Borzelli GLE, Geng T, Mackallah C, Ng B, Hong CC, Cai W, Huang  
 678 AY, Bodman R (2021) Generation of westerly wind bursts by forcing outside the tropics.  
 679 *Scientific Reports* 11(1):912, doi:10.1038/s41598-020-79655-7, URL [https://doi.org/10.  
 680 1038/s41598-020-79655-7](https://doi.org/10.1038/s41598-020-79655-7)
- 681 Takahashi K, Montecinos A, Goubanova K, Dewitte B (2011) ENSO regimes:  
 682 Reinterpreting the canonical and Modoki El Niño. *Geophysical Research Letters*  
 683 38(10), doi:<https://doi.org/10.1029/2011GL047364>, URL [https://doi.org/10.1029/  
 684 2011GL047364](https://doi.org/10.1029/2011GL047364)
- 685 Tippett MK, L’Heureux ML (2020) Low-dimensional representations of Niño 3.4 evolu-  
 686 tion and the spring persistence barrier. *npj Climate and Atmospheric Science* 3(1):1–11,  
 687 doi:10.1038/s41612-020-0128-y, URL <http://dx.doi.org/10.1038/s41612-020-0128-y>
- 688 Trenberth KE (2019) El Niño southern oscillation (ENSO). *Encyclopedia of Ocean Sciences*  
 689 (March):420–432, doi:10.1016/B978-0-12-409548-9.04082-3
- 690 Vimont DJ, Wallace JM, Battisti DS (2003) The seasonal footprinting mechanism in the  
 691 Pacific: Implications for ENSO. *Journal of Climate* 16(16):2668–2675, doi:10.1175/1520-  
 692 0442(2003)016<2668:TSMIT>2.0.CO;2
- 693 Wang C, Picaut J (2004) Understanding Enso Physics—A Review.  
 694 doi:<https://doi.org/10.1029/147GM02>, URL <https://doi.org/10.1029/147GM02>
- 695 Wyrтки K (1975) El Niño—The Dynamic Response of the Equatorial Pacific Ocean to At-  
 696 mospheric Forcing. *Journal of Physical Oceanography* 5(4):572–584, doi:10.1175/1520-  
 697 0485(1975)005<0572:ENTDRO>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
 698 journals/phoc/5/4/1520-0485\\_1975\\_005\\_0572\\_entdro\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/phoc/5/4/1520-0485_1975_005_0572_entdro_2_0_co_2.xml)
- 699 Wyrтки K (1985) Water displacements in the Pacific and the genesis of El  
 700 Niño cycles. *Journal of Geophysical Research: Oceans* 90(C4):7129–7132,  
 701 doi:<https://doi.org/10.1029/JC090iC04p07129>, URL [https://doi.org/10.1029/  
 702 JC090iC04p07129](https://doi.org/10.1029/JC090iC04p07129)
- 703 Zebiak SE (1989) Oceanic Heat Content Variability and El Niño Cy-  
 704 cles. *Journal of Physical Oceanography* 19(4):475–486, doi:10.1175/1520-  
 705 0485(1989)019<0475:OHCVAE>2.0.CO;2, URL [https://journals.ametsoc.org/  
 706 view/journals/phoc/19/4/1520-0485\\_1989\\_019\\_0475\\_ohcvae\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/phoc/19/4/1520-0485_1989_019_0475_ohcvae_2_0_co_2.xml)
- 707 Zebiak SE, Cane MA (1987) A Model El Niño–Southern Oscillation. *Monthly Weather*  
 708 *Review* 115(10):2262–2278, doi:10.1175/1520-0493(1987)115<2262:AMENO>2.0.CO;2,  
 709 URL [https://journals.ametsoc.org/view/journals/mwre/115/10/1520-0493\\_1987\\_  
 710 115\\_2262\\_amen\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/mwre/115/10/1520-0493_1987_115_2262_amen_2_0_co_2.xml)
- 711 Zhang C, Gottschalck J (2002) SST Anomalies of ENSO and the Madden–Julian Oscilla-  
 712 tion in the Equatorial Pacific. *Journal of Climate* 15(17):2429–2445, doi:10.1175/1520-  
 713 0442(2002)015<2429:SAOeat>2.0.CO;2, URL [https://journals.ametsoc.org/view/  
 714 journals/clim/15/17/1520-0442\\_2002\\_015\\_2429\\_saoeat\\_2\\_0\\_co\\_2.xml](https://journals.ametsoc.org/view/journals/clim/15/17/1520-0442_2002_015_2429_saoeat_2_0_co_2.xml)

## 715 Funding

716 This research was supported by grant #19-42-04121 from the Russian Science Foundation  
 717 (selecting the state variables and constructing the data-driven models of ENSO). Also, it  
 718 was supported by grant #19-02-00502 from the Russian Foundation for Basic Research  
 719 (studying the ENSO phase locking to the annual cycle).

---

720 **Author contributions statement**

721 The authors contributed equally to the methodology and analysis of the results. A.S. per-  
722 formed calculations and wrote the initial manuscript.

723 **Data Availability Statement**

724 The ERSST data set was downloaded from the NOAA National Centers for Environmental  
725 Information (Huang et al., 2017). The OHC data provided by the Institute of Atmospheric  
726 Physics Chinese Academy of Sciences are available at <https://pan.cstcloud.cn/s/sloceeVQjo>.  
727 The monthly CO<sub>2</sub> trends provided by the NOAA Global Monitoring Laboratory are available  
728 at <https://gml.noaa.gov/aftp/products/trends/co2/>.

729 **Conflict of interest**

730 The authors declare that they have no conflict of interest.