

¹ **Linear predictability: A sea surface height case study**

² Maike Sonnewald* and Carl Wunsch[†]

³ *Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139*

⁴ Patrick Heimbach

⁵ *The University of Texas at Austin, 201 East 24th Street, Austin, TX 78712, USA and*

⁶ *Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA*

⁷ *Corresponding author address: Maike Sonnewald, Massachusetts Institute of Technology, 77
⁸ Massachusetts Ave, Cambridge, MA 02139, USA.

⁹ E-mail: maike_s@mit.edu

¹⁰ [†]Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA and
¹¹ Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA.

ABSTRACT

12 A benchmark of linear predictability of sea surface height (SSH) globally
13 is presented, complementing more complicated studies of SSH predictability.
14 Twenty years of the Estimating the Circulation and CLimate of the Ocean
15 (ECCOv4) state estimate (1992-2011) are used, fitting autoregressive moving
16 average ($\text{ARMA}(n,m)$) models where the order of the coefficients is chosen
17 by the Akaike Information Criteria (AIC). Up to 50% of the ocean SSH
18 variability is dominated by the seasonal signal. The variance accounted
19 for by the non-seasonal SSH is particularly distinct in the Southern and
20 Pacific Ocean, containing $> 95\%$ of the total SSH variance and the expected
21 prediction error growth taking a few months to reach a threshold of 1 cm.
22 Isolated regions take twelve months or more to cross an accuracy threshold of
23 1 cm. Including the trend, significantly increases the time taken to reach the
24 threshold, particularly in the South Pacific. Annually averaging has expected
25 prediction error growth of a few years to reach a threshold of 1 cm. Including
26 the trend mainly increases the time taken to reach the threshold, but the
27 timeseries is short and noisy.

28

29 **1. Motivation**

30 The variability and change of future sea surface height (SSH, denoted η) is the center of much
31 of the concern about the ongoing global warming. Understanding, and predicting these key values,
32 globally and regionally, involves projection and space-time integration of the numerous factors
33 that influence it, including the wind field, atmospheric pressure, tides, ice-melt, river runoff, heat
34 and freshwater exchange, and the shifting ocean circulation itself (Parker 1991; Church *et al.*
35 2013). The diverse physics spans a large range of timescales for oceanic response (e.g., Wunsch
36 (2015)).

37

38 Compared to the atmosphere, most relevant oceanic timescales are very long-ranging from
39 months to thousands of years. The presence of that long time-scale (long memory), and the
40 observed small perturbations in the oceanic state suggests that major elements (not all) of future
41 η can be predicted from a knowledge of the present and past states of the ocean. The expected
42 prediction error (PE) growth of η is not well established. Attempting to estimate the PE via en-
43 sembles of climate model simulations reveals large ensemble spread. (e.g, the Intergovernmental
44 Panel on Climate Change, IPCC) (Church *et al.* 2013; Stainforth *et al.* 2005; Palmer 2012).

45

46 The goal here is to assess quantitatively the extent to which η variability is predictable using
47 linear methods, describing both the deterministic (seasonal changes) as well as the underlying
48 continuum treated as a wide-sense stationary linear process.¹ As discussed e.g., by Wunsch
49 (2013), such an approach provides a baseline against which predictions made with consider-
50 ably more complex methods (non-linear, non-stationary, extended to spatial structure) can be

¹“Wide-sense” stationarity is the terminology of electrical engineering; mathematicians call it “weakly” stationary, and in both cases only the first two moments (mean and variance) are assumed time-independent (Priestley 1981).

51 compared. The general case involves much more complex computations, and raises the purely
52 practical issue of whether the linear, univariate, stationary approach is adequate and for how long?

53

54 An extensive body of literature explores the variability of η with varying degrees of complica-
55 tion ranging from statistical to applying hierarchies of general circulation models. These methods
56 have varying degrees of regional success (Gille 1994; Chowdhury *et al.* 2007; Melillo *et al.* 2014)
57 or global application (Church *et al.* 2013; Rahmstorf *et al.* 2012). The purely statistical approach
58 is less concerned with capturing the underlying physics, while the general circulation model
59 (GCM) approach treats the η field as the integrated sum of ocean and atmospheric processes as
60 reflected by the present physics. This study uses the simplest statistical approach to present a
61 benchmark for more complex studies.

62

63 Treating oceanic change as linear may be counter-intuitive. However, the modern observational
64 record shows no major shifts in the large-scale baroclinic structure of the ocean (e.g., Roemmich
65 *et al.* (2012)). Apart from small regions of sea ice and convection, well-understood theory also
66 supports the inference of only perturbation changes over decades to centuries (Hirschi *et al.* 2013).

67

68 Interpretation of statistics from short records is difficult (see e.g., Wunsch (1999); Percival *et al.*
69 (2001); Ocaña *et al.* (2016)). The methods that underlie much of what is presented here rely on
70 the assumptions that η changes from the superposition of deterministic seasonal components and
71 a wide-sense stationary stochastic process. Of most relevance for the latter are general red-noise
72 processes and the extreme of white noise, which is by definition linearly unpredictable. Detection
73 of true non-stationarity is not possible with the short records at hand. Similarly, an infinite number
74 of generalizations to non-linear representations are possible, but unless the linear assumption can

75 be excluded, it becomes an important reference point.

76

77 Local and global predictability are in many ways distinct, with e.g., regional variability in
78 η having been attributed to shifts in wind features such as the Pacific Decadal Oscillation and
79 the North Atlantic Oscillation (Yin and Goddard 2013). Here, the approach is basically that
80 of a univariate “black-box”, with the underlying mechanisms, e.g., determining the changing
81 global mean of η , having been discussed in many published papers (Ocaña *et al.* 2016; Parker
82 1991; Piecuch and Ponte 2011; Forget and Ponte 2015). The oceans are storing large portions
83 of the additional heat from global warming, the land ice is retreating, along with other external
84 forcings, but disucssion of these specific physical contributions as functions of time and position
85 is postponed.

86

87 The methods are detailed in Section 2, and the results are presented in Section 3 where the sea-
88 sonal and non-seasonal contributions to the variance of η are presented. As defined, the seasonal
89 component is perfectly predictable, and the non-seasonal portion involves stochastic forecasting.
90 A set of four experiments is presented: using monthly and annual means of η , with apparent lin-
91 ear trends included as part of the background rednoise, and with the trends removed. Section 4
92 presents the discussion and conclusion.

93 2. Numerical and ARMA models

94 Predictability is studied using the ECCOv4 global bi-decadal state estimate, as described by
95 Wunsch and Heimbach (2013); Forget *et al.* (2015) and others. This estimate is obtained by
96 the use of least-squares with Lagrange multipliers. The result is an adjusted, yet *free-running*
97 version of the MIT General Circulation Model (MITgcm, Adcroft *et al.* (2004)). In contrast

98 to most “reanalysis” products, the oceanic state satisfies basic conservation laws for enthalpy,
99 salt, volume, and momentum, remaining largely within error estimates of a diverse set of global
100 data (Wunsch and Heimbach 2013; Stammer *et al.* 2016; Wunsch and Heimbach 2007). The
101 ECCOv4 state estimate is generally within the best available uncertainty estimates of ocean
102 observations, including altimetry. Regions without data are filled in a dynamically consistent way
103 using the adjoint dynamics, avoiding issues using pure altimetry as in Reynolds *et al.* (2013).
104 Thus ECCOv4 is better-suited to discussing the variability of η than ordinary unadjusted forward
105 models, altimetry only, or “reanalysis” type products that do not satisfy basic conservation laws.

106

107 At each point of latitude and longitude (θ, λ) , the mean is removed and $\eta(\theta, \lambda)$ is defined
108 as $\eta = \eta'(\theta, \lambda) + \bar{\eta}(\theta, \lambda)$, where $\bar{\eta}(\theta, \lambda)$ denotes the seasonal and $\eta'(\theta, \lambda)$ the non-seasonal
109 $\eta(\theta, \lambda)$. Throughout this study, each θ, λ point is used and for simplicity the spatial indicies are
110 dropped hereafter. At this stage, any trend is being included as part of η' .

111

112 The seasonal component includes its first two harmonics as illustrated in Figure 1. Variances of
113 the seasonal and residual components are additive:

$$\sigma_{\eta}^2 = \sigma_{\eta'}^2 + \sigma_{\bar{\eta}}^2. \quad (1)$$

114 To the extent that $\sigma_{\bar{\eta}}^2 \gg \sigma_{\eta'}^2$, useful prediction is purely deterministic. When the annual
115 variability is not dominant, examination of the predictability of the residual process has to be
116 examined, and these two approaches are treated separately. Deterministic prediction of sinusoidal
117 components is straightforward.

118

119 Linear predictability of wide-sense stationary stochastic processes, not distinguishable from
 120 Gaussian, is well-understood with a very large literature including standard textbooks (e.g., Box
 121 and Jenkins (1970)). Here the formalism is discussed only insofar as it develops the notation to
 122 be applied. Most linear methods are based on the autoregressive process of order n (AR(n)), the
 123 moving average process of order m , (MA(m)), or the mixed autoregressive moving average of
 124 order n, m (ARMA(n, m)).
 125

126 As the textbooks show (e.g., Box and Jenkins (1970)), these representations are interchangeable,
 127 with a choice being mainly one of convenience, efficiency of representation or a combination of
 128 those two. If $\zeta(t)$ is a zero-mean wide-sense stationary time-series, two of the representations are,

$$\text{AR}(n) : \quad \zeta(t) = \underbrace{a_1}_{\text{coeff.}} \zeta(t-1) + a_2 \zeta(t-2) + \dots + a_n \zeta(t-n) + \overbrace{e(t)}^{\text{white noise}}, \quad (2)$$

$$\text{MA}(m) : \quad \zeta(t) = e(t) + \underbrace{b_1}_{\text{coeff.}} e(t-1) + b_2 e(t-2) + \dots + b_m e(t-m). \quad (3)$$

129 A combination gives the general ARMA(n, m) model:

$$\zeta(t) = a_1 \zeta(t-1) + a_2 \zeta(t-2) + \dots + a_n \zeta(t-n) + e(t) + b_1 e(t-1) + b_2 e(t-2) + \dots + b_m e(t-m), \quad (4)$$

130 where a_i and b_i are regression coefficients, $e(t)$ is near-Gaussian white noise with zero mean and
 131 variance σ_e^2 , and t is any time, past, present or future, measured in units producing an implied
 132 time-step $\Delta t = 1$. Conversion of one form to another, or to the mixed representation is discussed
 133 in textbooks (e.g., Box and Jenkins (1970)). In the MA form, the white noise increments $e(\tau)$
 134 are determined for past values leading up to the present time, t . $e(t)$ is known (estimated),
 135 but no future values $e(t + \Delta t)$, $\Delta t > 0$ are known. In the AR form, enough past values $\zeta(\tau)$
 136 are assumed known, as is $e(t)$, but again no future values are available. In the presence of

¹³⁷ noise, these representations can become unstable, being indistinguishable from the presence
¹³⁸ of apparent non-stationarity. Tests for stability/non-stationarity are based upon the zeros and
¹³⁹ poles of complex polynomials formed from the various coefficients, a_i, b_j (Box and Jenkins 1970).

¹⁴⁰

¹⁴¹ For numerical efficiency, the relation AR(1)=MA(∞) is used to estimate PE. Converting the
¹⁴² ARMA to the MA form using the Wold representation the τ -ahead PE is:

$$\langle (\hat{\zeta}(t+\tau) - \zeta(t+\tau))^2 \rangle = \underbrace{\sigma_\varepsilon^2}_{\text{noise variance}} \sum_{p=0}^{\tau} \underbrace{b_p^2}_{\text{coeff.}}, b_0 = 1. \quad (5)$$

¹⁴³ This quadratic error growth depends upon the values of b_p , whose sum can never exceed the time
¹⁴⁴ series variance $\langle \zeta^2 \rangle$.

¹⁴⁵

¹⁴⁶ The performance of the ARMA(n, m) is assessed in terms of the relatively small PE growth over
¹⁴⁷ time. This criterion is expressed as the time it takes the error to grow beyond a given threshold,
¹⁴⁸ and good performance refers to a small PE at a specific time.

¹⁴⁹

¹⁵⁰ In practice, regression coefficients a_i, b_j are most-often found using one of several versions of
¹⁵¹ least-squares in which autocovariances are estimated along the way. The most common difficulty
¹⁵² is determining the orders, n, m for the particular representation, with the orders being increased
¹⁵³ incrementally by one until a stopping criterion is met. (See for example, Hughes and Williams
¹⁵⁴ (2010), Aho *et al.* (2014) or Akaike (1973)). Adding regression parameters improves the fit to
¹⁵⁵ the data, but risks over-fitting. The Akaike Information Criteria (AIC) is used to minimize the
¹⁵⁶ expectation of the PE where k is the number of parameters:

$$\text{AIC} = 2k - 2\ln(\mathcal{L}),$$

¹⁵⁷ where \mathcal{L} is the likelihood:

$$\mathcal{L} = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\sum_{i=1}^N \frac{(\zeta_i - \hat{\zeta}_i)^2}{2\sigma^2}\right).$$

¹⁵⁸ ζ_i is the observed, and $\hat{\zeta}_i$ is the prediction, so $(\zeta_i - \hat{\zeta}_i)^2$ are the prediction residuals. In this manner
¹⁵⁹ the AIC value is minimized, which determines the smallest appropriate order to represent the
¹⁶⁰ time-series. As discussed e.g., by Priestley (1981) and Yang (2005), the AIC can overestimate the
¹⁶¹ order, see Appendix A1.

¹⁶²

¹⁶³ In the following, we consider time series where only a time mean has been removed (η)
¹⁶⁴ and those where a linear trend has been removed as well (η^\dagger) for analyses of monthly and
¹⁶⁵ annual time series. Further removing a seasonal cycle $\bar{\eta}$ leads to time series η' and η'^\dagger ,
¹⁶⁶ respectively. A significant linear or quadratic trend can itself be used to make a prediction. By
¹⁶⁷ including that structure in the stochastic process, the intrinsic predictability of the process may
¹⁶⁸ be over-predicted, but it accounts in large part for the true predictability including that of any trend.

¹⁶⁹

¹⁷⁰ 3. Results

¹⁷¹ a. Seasonal Variance

¹⁷² Figure 3 shows the total variance σ_η^2 of η from 1992-2011. Western boundary currents and
¹⁷³ their extensions are associated with higher variance, particularly in the Northern Hemisphere.
¹⁷⁴ The tropics have a large variance, particularly in the Pacific Ocean with the Indian Ocean being

175 dominated by the Arabian Sea monsoonal effects. Around 15°S in the Indian Ocean there is a
176 streak of higher variability tilted southward. Notable elevated standard deviations occur along
177 the eastern side of the Indian Ocean basin, that are not seen in the Atlantic or Pacific Oceans.
178 The Southern Ocean shows some excess variance, particularly in the Indian and Pacific sectors.
179 Bathymetric features such as the Pacific-Antarctic ridge are associated with higher variance
180 (Ponte and Piecuch 2014).

181

182 Figure 4 illustrates the percentage of the total variance included in the seasonal η component. A
183 striking, but well-known, seasonal hemispheric difference appears, where much of the η variance
184 in the Northern Hemisphere is dominated by the seasonal component. Exceptions include a large
185 zonal streak in the Pacific and areas at a similar latitude in the North Atlantic. The Irminger and
186 Labrador Seas also have large areas where the seasonal signal is less dominant, as is also true of
187 the Bering Sea. The South Pacific Ocean has large areas where the seasonal signal is weak, but a
188 strong patch extends westwards off the coast of Peru, likely associated with the upwelling there,
189 as well as the East Australian Current. In the Indian Ocean, fluctuations in the Arabian Sea are
190 almost entirely captured by the seasonal component. The eastern half of the Indian Ocean in the
191 Southern Hemisphere is largely dominated by $\bar{\eta}$, but less clearly so in the East. In the Southern
192 Ocean a clear seasonal dominance appears in the Weddell Gyre region, as well as in the area of
193 the standing meander off the coast of South Africa.

194

195 *b. Seasonal Prediction*

196 The seasonal signal, $\bar{\eta}$, is predictable as defined over the bi-decadal time interval covered
197 by ECCOv4, and Figure 5 illustrates the associated contribution to the prediction in cm over

198 1992-2011 taken as the total contribution to the variance. Generally speaking, where the seasonal
199 variance is a large fraction of the total, good predictability is found, e.g associated with the
200 seasonal component in the Mascarene basin area, as well as in the Southern Ocean in the Indian
201 and Pacific Ocean sectors.

202

203 *c. Non-seasonal Variance*

204 Figure 6 shows the percentage of the total variance accounted for by the residual background
205 process. Here a striking signal appears across the Pacific Ocean along the equator, likely
206 associated with the El Niño Southern Oscillation (ENSO) climate mode. Further zonal streaks
207 follow to the north. A strong signal occurs in the Bering Strait region. In the Atlantic Ocean, the
208 subtropical regions show active areas, and along the paths of the Labrador and southern tip of the
209 East Greenland currents. In the South Atlantic, the non-seasonal component of variance accounts
210 for a large portion of the variance in the Brazil current and Zapiola region, as well as in zonal
211 streaks. In the Indian Ocean, the non-seasonal component of variance also accounts for a large
212 portion of the total variance to the west in the Mascarene basin region, as well as along the western
213 Indonesian coast. The Southern Ocean produces a very large signal associated with regions of
214 deep mixed layers and possibly mode water formation and areas where the ACC heads south.
215 Predictability associated with the non-seasonal variance is addressed in the remainder of this paper.

216

217 *d. Predictability After Trend Removal*

218 A linear trend is now removed, meaning that a possibly perfectly predictable component
219 is eliminated. The η^{\dagger} is fit to an ARMA(n, m) process of the ECCOv4 state estimate from

1992-2011. A four point smoother is applied to reduce noise (two points in latitude and two in longitude), equivalent to moving from the tracer (t -point) to the vorticity point (f -point) in the Arakawa C-grid. Using the smoother tends to make the data adhere more closely to a normal distribution, but can exaggerate the spatial covariance of isolated outliers. Figure 2 illustrates the effect of different order choices of n and m . The performance of the different choices is given by the rate of the PE growth over time. Figure 7a show the order n of the ARMA(n, m) chosen with the AIC. The order chosen shows how many coefficients are used to optimally represent η'^\dagger and the associated prediction error.

228

The simplest linear theory assumes that the underlying data are Gaussian or close to it, an assumption tested in Figure 8a using the Shapiro-Wilk test for normality (Shapiro and Wilk 1965). Large areas associated with features such as the ENSO signal appear to deviate from normality (i.e., p values close to 0). This result has implications for the predictability since these departures are important when interpreting the PE.

234

Figure 9 shows the error growth asymptoting to its upper-bound: the full variance of the background residual time series. Large differences as a function of region appear in the PE, as well as in their asymptotic rate of growth. The expected error e -folding structure is shown to illustrate the rate of predictability decay, independent of magnitude.

239

The ARMA(n, m) expected PE growth associated with the η'^\dagger is illustrated in figures 10a. The PE growth is expressed in terms of the time taken (months) before a target is exceeded (1cm). Large areas of the ocean show limited performance based on the PE growth over time, but certain (mostly isolated) areas have good performance, exceeding a year. In interpreting

244 these figures, it is prudent to keep the assumption of normality in mind (Figure 8a). Figure 10a
245 shows the ARMA(n, m) expected PE growth to a 1 cm threshold. High predictability is seen in
246 the North Pacific Ocean. Although beyond the intended present scope of this paper, these areas
247 are associated with the Kuroshio crossing the Pacific and wave-eddy interactions stretching from
248 Hawaii to the coast of California in a banana-shape. Areas in the Equatorial Pacific also show
249 better performance of PE growth over time, but are clearly non-normal. Features associated with
250 the Pacific-Antarctic Ridge in the Southern Ocean also produce better performance in terms of
251 the ARMA(n, m) expected PE growth over time, along with some areas in the Irminger Sea and a
252 streak in the Arabian Sea.

253

254 *e. Predictability with Apparent Trends*

255 Tests of predictability are now made with the linear trend left in the time series. Including
256 the trend treats it as an unresolved component of a rednoise process. The inclusion of the trend
257 is expected to increase the performance of the ARMA(n, m) PE growth over time. One cannot
258 distinguish this variability from a rednoise process with existing data (Church *et al.* 2013; Lyu *et*
259 *al.* 2014; Ocaña *et al.* 2016). To assess its impact, the background analysis is repeated, including
260 the trend. Figures 7c and 7d illustrates the ARMA(n, m) order chosen by the AIC. This result is
261 similar to Figures 7a and 7b, but sometimes smaller n, m result as is physically plausible.

262

263 The normality of the background process with the trend is re-tested in Figure 8b, illustrating
264 that most of the ocean remains indistinguishable from a normal distribution in η' . Figure
265 10b shows the prediction performance based on the ARMA(n, m), phrasing the result in terms
266 of the number of months it takes for the PE to cross the accuracy target of 1 cm. Retaining

267 the trend adds predictability, with large areas exceeding 12 months before exceeding the 1
268 cm threshold. Areas where the trend is important are in bands in the subtropics in both the
269 Atlantic and Pacific, as well as large areas off the coast of Greenland, the Drake passage, the
270 Kuroshio path across the Pacific, the southern Indian Ocean and a large region in the South Pacific.

271

272 *f. Predictability with Annual Averages*

273 Interannual and monthly physics are distinct. Assessing the annually averaged η'^\dagger and η' from
274 ECCOv4 separately, the assumptions of estimating the covariance and of a Gaussian distribution
275 are likely inaccurate, owing to the short, 20-datapoint record.

276

277 Figure 13 shows the chosen ARMA(n, m) order. A four-point smoother is again used for re-
278 duced noise on the ECCOv4 η'^\dagger 1992-2011 where the trend is removed and the data are annually
279 averaged. The n and m of the ARMA(n, m) now reflect the annually averaged data, rather than
280 the monthly. The n of the ARMA(n, m) is generally lower for the annually averaged η'^\dagger than for
281 monthly η'^\dagger . However, large regions show very different patterns than from the monthly data. Ex-
282 amples include the Arabian Sea region and the North Pacific. However, the areas where one order
283 dominates are generally larger than for monthly η'^\dagger . Regions that have higher orders for yearly
284 averaged data than for monthly averages are areas such as the Pacific sector of the Southern Ocean.

285

286 Figure 11a assesses the extent to which the 20-year time series can be viewed as coming from a
287 normally distributed population when using annual averages. Most of the ocean passes this test,
288 but with large areas that appear noisy/non-stationary, presumably owing primarily to the presence

289 of transient noise structures.

290

291 The associated prediction performance based on the ARMA(n, m) expected PE growth over
292 time is shown in Figure 12a. The physics that lead to annual predictability are different from
293 the monthly; thus the large change in spatial patterns is expected. In some areas the time to
294 reach the criterion of 1 cm expected error is several years. These regions are clustered in the
295 Pacific, with a patch in the western equatorial Pacific, a band stretching from the central North
296 Pacific to the coast of the USA (Hawaii to California), as well as isolated small patches elsewhere.
297 A larger patch of good PE performance over time exists in the Pacific sector of the Southern Ocean.

298

299 As expected, the role of the linear trend is likely also important in the annually averaged η' .
300 Figure 13 illustrates the associated orders of the ARMA(n, m). As with the monthly data, the
301 orders are generally lower. Higher orders exist in a band stretching from the central Pacific north-
302 wards to the USA (Hawaii to California) as well as a feature associated with the Antarctic-Pacific
303 ridge.

304

305 Figure 11b suggests that most of the ocean passes the test of normality, but again with large
306 areas of failure. The associated prediction performance based on the ARMA(n, m) expected PE
307 growth over time is shown in Figure 12b. This result is similar to that in Figure 12a, but most
308 of the longer-term PE performance over time is found in the South Pacific. A band stretching
309 from the central Pacific northwards to the USA still exists, but large areas do not pass the test of a
310 stable or wide-sense stationary ARMA. Patches of longer term PE performance over time are also
311 seen in the Indian Ocean and Drake Passage, with isolated regions elsewhere.

312

313 **4. Discussion and Conclusions**

314 In this paper, linear univariate predictability of sea level, η , is discussed, and benchmarks are
315 presented. In general, more complex schemes would need to exceed this PE performance over
316 time to be proven worthwhile. Prediction performance is presented in terms of the time it takes the
317 expected prediction error (PE) to grow beyond 1 cm. More complex models should necessarily
318 do better and may well be justified, particularly in specific physically identifiable regions. This
319 approach is supported by work such as Goddard *et al.* (2015), where certain events in η' have
320 been attributed to factors such as changes in the Atlantic meridional overturning circulation and
321 the North Atlantic Oscillation. Existing spreads in ensemble studies, such as the CMIP5 models,
322 would suggest that many difficulties remain (Church *et al.* 2013).

323

324 In the ECCOv4 state estimate, the seasonal cycle, $\bar{\eta}$, accounts for more than 80% of the
325 variability over 20+ years in large parts of the Atlantic and Northern Hemisphere Pacific, as well
326 as the Weddell Gyre area in the Southern Ocean. In these regions, the seasonal component is
327 likely sufficient for estimating the variation in η for at least a few decades.

328

329 The percentage of 20-year variance in the stochastic residual, η' , is important over large areas of
330 the ocean, particularly in the Southern Hemisphere. η' is treated as a weakly stationary univariate
331 random process, with the assumption of being normally distributed. Areas where η' is particularly
332 important for predictions lie in the Southern Ocean, and in extensive regions of the Pacific as well
333 as in the Indian Oceans. When a linear trend is removed, regions where the ARMA prediction
334 still performs well are mainly clustered in the Pacific, particularly in a band extending northeast
335 from Hawaii, as well as in isolated small areas elsewhere. A patch of higher predictability exists

336 in the Pacific sector of the Southern Ocean. However, the ARMA prediction, as expected, tends
337 to perform less well when the trend is removed.

338

339 An important next step is to distinguish the physical mechanisms, whether atmospheric or
340 oceanic, where the ARMA prediction procedure does well and those where it works poorly.
341 Investigating the contribution of specific mechanisms to the predictability of η is outside the
342 scope of this study, but results presented suggest such analysis is merited. For example, work
343 with linear models of η , includes that of Hughes and Williams (2010). Using altimetry alone,
344 and without a prediction focus, they fit AR(n) models, choosing the orders using the Bayesian
345 Information Criterion (BIC). With their higher temporal resolution, the AR(n) fits are not easily
346 compared to those from the use of the ARMA(n, m) on monthly and annual average values.
347 They concluded that Rossby wave patterns are important within +/-30° of the equator, while
348 advective processes become more influential at higher latitudes allowing features such as the
349 Pacific-Antarctic ridge to influence η' predictability. The ARMA fits to the ECCOv4 data also
350 have distinct features associated with the Pacific-Antarctic ridge and different behavior within
351 30° of the equator, suggesting the physics that give rise to these features are coherent across these
352 timescales and lend themselves to linear modeling approaches.

353

354 Predictability from annually averaged data proves generally different. With the trend included,
355 a region in the western South Pacific Ocean has striking performance. Removing the trend, an
356 increase in prediction performance is seen in a band extending north-east of Hawaii, likely due to
357 noise. Given the long time scales controlling oceanic physics, the records remain far too short to
358 infer statistically stable results. In this context, the continuing difficulties generally experienced
359 in distinguishing the lowest frequencies present between a general rednoise process and a true

secular trend of multi-decadal applicability, remains a major issue. Whether unconstrained models, such as the CMIP5 ones used by Lyu *et al.* (2014), have true prediction skill remains unknown. Note too, that the univariate approach used here is readily extended to accommodate multivariate predictive models employing correlated spatial structures of many different types and which may work much more effectively in some areas.

365

In brief summary: the present study produces a benchmark of univariate linear skill in predicting η . Up to 50% of the ocean is predictable using only the seasonal signal. The remaining ocean has an expected prediction error growth taking a few months to exceed 1 cm. Treating the linear trend as part of the continuum enhances the predictive performance as expected. In areas in the Southern and Pacific Oceans, the stochastic continuum, η' , contains more than 95% of the total variance of η , with expected prediction performance of 1 cm exceeding a year in significant portions of these regions. Moving forwards, extending the global measurements of η and understanding the underlying physical processes remain the key to progress.

374 **5. Acknowledgments**

375 This work was funded by the US National Aeronautics and Space Administration Sea Level
376 Change Team (contract NNX14AJ51G) and through the ECCO Consortium funding via the Jet
377 Propulsion Laboratory.

378 **APPENDIX**

379 **Influence of chosen Information Criteria**

380 **A1. Influence of chosen Information Criteria**

381 The choice of information criteria to determine the ARMA(n, m) significantly influences the
382 predictive performance. The AIC (equation 2) was used, as it demonstrated better performance for
383 applications where the "true" model is likely not available (Yang 2005). The Bayesian Information
384 Criteria (BIC) has been used in studies such as Hughes and Williams (2010). The BIC is defined
385 as:

$$\text{BIC} = k \ln(n) - 2 \ln(\mathcal{L}),$$

386 where n is the number of datapoints, k the number of parameters and \mathcal{L} is the likelihood shown
387 in equation 2.

388

389 As discussed by Priestley (1981) and Yang (2005) the AIC tends to overestimate the true order,
390 and the BIC to underestimate it. The AIC results showed better predictive power. The two criteria
391 give different weights to penalizing the number of regression coefficients, with the BIC having a
392 larger penalty term.

393

394 Figure 14 illustrates the orders chosen by the BIC for η'^\dagger . Note the difference to Figure 7 where
395 higher orders are chosen using the AIC. Overall, the AIC produces estimates at higher orders than
396 the BIC, but results for large areas are similar: e.g. the North Pacific and equatorial Atlantic. The
397 information criteria, particularly the AIC, tend to pick out differing dynamical regions.

398

399 The predictive potential associated with the η'^\dagger using the BIC is illustrated in Figure 15. This
400 shows a very similar pattern to Figure 10a, and we highlight the differences by showing the
401 difference between the two (AIC-BIC) in Figure 16a. Here we see that the AIC largely gives
402 higher predictive performance based on the ARMA(n, m) expected prediction error growth. The
403 difference between the AIC and BIC prediction performance is highlighted in Figure 16b. We
404 see slight differences, with the AIC having higher predictive performance to in bands $+/-30^\circ$ of
405 the Equator, particularly in the Pacific and Indian Ocean but also in the Atlantic. The BIC shows
406 somewhat higher predictive performance in the higher latitudes, particularly in the Pacific.

407

408 For annually averaged η' the difference between using the AIC and BIC are smaller. Figure 16c
409 highlights the difference between the AIC and BIC criterion, showing that overall the AIC has
410 higher performance based on the ARMA(n, m) expected prediction error growth. With annually
411 averaged η'^\dagger , the AIC and BIC also give similar prediction accuracies, with the differences
412 highlighted in the difference plot shown in Figure 16d. However, small areas show higher
413 performance based on ARMA(n, m) expected prediction error growth with the BIC.

414

415 The difference in predictability with the AIC and BIC informing the choice of the order suggests
416 that the AIC has the highest utility. The BIC underestimates the order, and the AIC is found to be
417 more suitable. This is illustrated in detail looking at the predictability with the apparent trend in η'

418 in section d, and demonstrated throughout with the higher prediction performance using the AIC.
419 The AIC is better particularly in the equatorial regions. The difference between the performance
420 of the prediction is larger in the monthly data, with the trend. For the case that is not detrend,
421 the BIC occasionally does better than the AIC, but no obvious spatial pattern is apparent. The
422 differences are larger for the annually averaged data, as the chosen ARMA(n, m) orders would
423 suggest. The higher orders chosen by the AIC is not surprising, as the AIC penalizes adding
424 parameters less strongly than the BIC. Burnham and Anderson (2002) show that the AIC can
425 actually be derived from the BIC using a different prior in the Bayesian framework. They suggest
426 the AIC has advantages over the BIC, firstly being based on information theory and secondly
427 having a more sensible prior. Our results are in agreement with Burnham and Anderson (2002),
428 and similar work by Yang (2005), suggesting better performance based on the ARMA(n, m).
429 expected prediction error growth is achieved using the AIC. The AIC has been seen to have
430 higher performance than the BIC as is discussed further by Burnham and Anderson (2004); Aho
431 *et al.* (2014). The implications of how well the different models capture the different dynamical
432 regimes is not discussed, as this relies on large generalizations of the prediction performance of
433 the fitted ARMA(n, m) over vast areas.

434

435 **References**

- 436 Adcroft, A., Hill, C., Campin, J. M., Marshall, J., & Heimbach, P. (2004). Overview of the formulation and numerics of the MIT GCM (pp.
437 139-150). Presented at the ECMWF Conference Proceedings, Shinfield Park, Reading, UK.
- 438 Aho, K.; Derryberry, D.; Peterson, T. (2014), "Model selection for ecologists: the worldviews of AIC and BIC", Ecology 95: 631-636,
439 doi:10.1890/13-1452.1.

- 440 Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N. ; Csiki, F., 2nd International
441 Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akademiai Kiad, p. 267-281.
- 442 Bingham, R. J., C. W. Hughes, V. Roussenov, and R. G. Williams (2007), Meridional coherence of the North Atlantic meridional overturning
443 circulation, *Geophys. Res. Lett.*, 34, L23606, doi:10.1029/2007GL031731.
- 444 Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- 445 Burnham, K. P.; Anderson, D. R. (2004), "Multimodel inference: understanding AIC and BIC in Model Selection" (PDF), *Sociological Methods*
446 & Research 33: 261-304, doi:10.1177/0049124104268644.
- 447 Burnham, K. P.; Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.),
448 Springer-Verlag, ISBN 0-387-95364-7.
- 449 Chelton, D. B., and M. G. Schlax, 1996: Global observations of oceanic Rossby waves. *Science*, vol. 272, pp. 234-238.
- 450 Chowdhury, M.D. R., Chu, P.-S., Schroeder, T. and Colasacco, N. (2007), Seasonal sea-level forecasts by canonical correlation analysisan opera-
451 tional scheme for the U.S.-affiliated Pacific Islands. *Int. J. Climatol.*, 27: 1389-1402. doi:10.1002/joc.1474
- 452 Cipollini, P., D. Cromwell, P. G. Challenor and S. Raffaglio, Rossby waves detected in global ocean colour data, *Geophysical Research Letters*,
453 Vol. 28, No. 2 , pp. 323-326, 2001.
- 454 Forget, G., J.-M. Campin, P. Heimbach, C. N. Hill, R. M Ponte, and C. Wunsch, ECCO version 4: an integrated framework for non-linear inverse
455 modeling and global ocean state estimation, *Geo. Sci. Model Dev.*, 8, 2015
- 456 Forget, G. and Ponte, R.M., The partition of regional sea level variability, *Progress in Oceanography*, 137, 173-195, 2015.
- 457 Gille, S. T., 1994. Mean sea surface height of the Antarctic Circumpolar Current from Geosat data: Method and application, *J. Geophys. Res.*, 99,
458 18,255-18,273.
- 459 Paul B. Goddard, Jianjun Yin, Stephen M. Griffies, Shaoqing Zhang. An extreme event of sea-level rise along the Northeast coast of North America
460 in 2009?2010. *Nature Communications*, 2015; 6: 6346 DOI: 10.1038/ncomms7346

- 461 Hare, S. R., and N. J. Mantua, 2000: Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography*, Vol. 47,
462 Pergamon, 103-146.
- 463 Hirschi, J. J.-M., Blaker, A. T., Sinha, B., Coward, A., de Cuevas, B., Alderson, S., and Madec, G.: Chaotic variability of the meridional overturning
464 circulation on subannual to interannual timescales, *Ocean Sci.*, 9, 805-823, doi:10.5194/os-9-805-2013, 2013.
- 465 Chris W. Hughes, Simon D. P. Williams, The color of sea level: Importance of spatial variations in spectral shape for assessing the significance of
466 trends, *Journal of Geophysical Research*, 2010, 115, C10
- 467 IPCC, 2013: Climate Change 2013: Church, J.A., et al., 2013: Sea Level Change. In: Climate Change 2013: The Physical Science Basis.
468 Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin,
469 G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Chapter 13, pp. 1137-1216, Cambridge
470 University Press.
- 471 Lin, X., J. Yang, D. Wu, and P. Zhai (2008), Explaining the global distribution of peak-spectrum variability of sea surface height, *Geophys. Res.*
472 Lett., 35, L14602, doi:10.1029/2008GL034312.
- 473 Lorenz, E.N.: Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, 20 (2), 130-141, doi: 10.1175/1520-0469(1963)
474 020<0130:DNF>2.0.CO;2, 1963.
- 475 Lorenz, E.N.: Low-order models and their uses, in G. N. Cathy Nicolis (ed.), *Irreversible phenomena and dynamical systems analysis in geo-*
476 *sciences*, D. Reidel Publishing Company, pp. 557-567, 1987.
- 477 Lyu, K., Zhang, X., Church, J.A., Slanen, A.B.A., Hu, J., (2014). Time of emergence for regional sea-level change. *Nature Climate Change* 4,
478 1006-1010. doi:10.1038/nclimate2397
- 479 J. V. Mecking, S. S. Drijfhout, L. C. Jackson, T. Graham (2016) Stable AMOC off state in an eddy-permitting coupled climate model. *Climate*
480 *Dynamics*
- 481 Mengel, M., Levermann, A. (2014): Ice plug prevents irreversible discharge from East Antarctica. *Nature Climate Change* (online) [DOI:
482 10.1038/NCLIMATE2226]

- 483 Melillo, J.M., T.C. Richmond and G.W. Yohe, Eds. (2014), Climate Change Impacts in the United States: The Third National Climate Assessment,
484 U.S. Global Change Research Program, 841 pp., doi:10.7930/J0Z31WJ2
- 485 Milly, Paul CD, Kathryn A. Dunne, and Aldo V. Vecchia. "Global pattern of trends in streamflow and water availability in a changing climate."
486 Nature 438.7066 (2005): 347-350.
- 487 Minobe, S., 1999: Resonance in bidecadal and pentadecadal climate oscillations over the North Pacific: Role in climate regime shifts. Geophys.
488 Res. Lett., 26, 855-858.
- 489 Ocaña, V., E. Zorita, and P. Heimbach (2016), Stochastic secular trends in sea level rise, J. Geophys. Res. Oceans, 121, 2183-2202,
490 doi:10.1002/2015JC011301.
- 491 Overland, J. E., J. M. Adams, and N. A. Bond, 1999: Decadal variability of the Aleutian low and its relation to high-latitude circulation. J. Climate,
492 12, 1542-1548.
- 493 Overland, J. E., J. M. Adams, and H. O. Mofjeld, 2000: Chaos in the North Pacific: Spatial modes and temporal irregularity. Progress in Oceanog-
494 raphy, Vol. 47, Pergamon, 337-354.
- 495 Palmer, T.N., (2012) Towards the probabilistic Earth-system simulator: a vision for the future of climate and weather prediction. Quarterly Journal
496 of the Royal Meteorological Society 138, 841-861.
- 497 "Sea Level as an Indicator of Climate and Global Change". Dr. Bruce Parker. Marine Technology Society Journal, Vol. 25, No. 4, 1991.
- 498 Donald B. Percival, James E. Overland, and Harold O. Mofjeld (2001) Interpretation of North Pacific Variability as a Short- and Long-Memory
499 Process, Journal of Climate, 14:24, 4545-4559
- 500 Piecuch, C. G., and R. M. Ponte (2011), Mechanisms of interannual steric sea level variability, Geophys. Res. Lett., 38, L15605,
501 doi:10.1029/2011GL048440.
- 502 Pierce, D. W., 2001: Distinguishing coupled ocean-atmosphere interactions from background noise in the North Pacific. Progress in Oceanography,
503 Pergamon, 49(1-4):331-352

- 504 Ponte, R.M. and C.G. Piecuch, 2014: Interannual Bottom Pressure Signals in the AustralianAntarctic and Bellingshausen Basins. *J. Phys.*
505 *Oceanogr.*, 44, 14561465, <https://doi.org/10.1175/JPO-D-13-0223.1>
- 506 Priestley, M. B. (1981) Spectral Analysis and Time Series, London: Academic Press.
- 507 Rahmstorf, S., M. Perrette, and M. Vermeer (2012), Testing the robustness of semi-empirical sea level projections, *Climate Dynamics*, 39(3-4),
508 861-875, doi: <http://dx.doi.org/10.1007/s00382-011-1226-7>.
- 509 Reynolds, R.W., D.B. Chelton, J. Roberts-Jones, M.J. Martin, D. Menemenlis, and C.J. Merchant, 2013: Objective Determination of Feature
510 Resolution in Two Sea Surface Temperature Analyses. *J. Climate*, 26, 2514-2533, <https://doi.org/10.1175/JCLI-D-12-00787.1>
- 511 Roemmich, D, Gould WJ, Gilson J. 2012. 135 years of global ocean warming between the Challenger expedition and the Argo Programme. *Nature*
512 *Climate Change*. 2:425-428. 10.1038/nclimate1461
- 513 Thomson, D.J. 1994. Jackknifing multiple-window spectra.In: Pro-ceedings of the IEEE International Conference on Acoustics, Speech and Signal
514 Processing, VI, 73-76.
- 515 Tulloch, R., J. Marshall, and K. S. Smith (2009), Interpretation of the propagation of surface altimetric observations in terms of planetary waves
516 and geostrophic turbulence, *J. Geophys. Res.*, 114, C02005, doi:10.1029/2008JC005055.
- 517 Shapiro, S.S. and Wilk, M.B. (1965). "An analysis of variance test for normality (complete samples)". *Biometrika*. 52 (3-4): 591-611.
518 doi:10.1093/biomet/52.3-4.591. JSTOR 2333709. MR 205384. p. 593.
- 519 Stainforth, D., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D., Kettleborough, J., Knight, S., Martin, A., Murphy, J., Piani, C., Sexton,
520 D., Smith, L., Spicer, R., Thorpe, A. and Allen, M.: 2005, Uncertainty in predictions of the climate response to rising levels of greenhouse
521 gases, *Nature* 433(7024), 403-406. url: <http://oro.open.ac.uk/5025/>
- 522 D. Stammer, M. Balmaseda, P. Heimbach, A.Koehl, and A. Weaver, 2016: Ocean Data Assimilation in Support of Climate Applications: Status
523 and Perspectives. *Ann. Rev. Mar. Sci.*, 8, 491-518.
- 524 Yang, Y. (2005), "Can the strengths of AIC and BIC be shared?", *Biometrika*, 92: 937-950, doi:10.1093/biomet/92.4.937.

- 525 Yin, J., and P. B. Goddard (2013), Oceanic control of sea level rise patterns along the East Coast of the United States, *Geophys. Res. Lett.*, 40,
- 526 5514?5520, doi:10.1002/2013GL057992.
- 527 Warren, Bruce A., Thomas Whitworth, and Joseph H. LaCasce. "Forced resonant undulation in the deep Mascarene Basin." *Deep Sea Research*
- 528 Part II: *Topical Studies in Oceanography* 49.7 (2002): 1513-1526.
- 529 Wunsch, C., and P. Heimbach, 2007: Practical global oceanic state estimation. *Physica D*, 230, 197-208.
- 530 Wunsch, C. and P. Heimbach, 2013, Dynamically and kinematically consistent global ocean circulation and ice state estimates. In *Ocean Circulation*
- 531 and Climate 2nd Edition, Siedler et al., Eds.
- 532 Wunsch, C. The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull. Am. Met. Soc.* 80,
- 533 245-255 (1999).
- 534 Wunsch, C. (2013), Covariances and linear predictability of the Atlantic Ocean, *Deep Sea Res., Part II*, 85, 228-243.
- 535 Wunsch, C. (2015), *Modern Observational Physical Oceanography: Understanding the Global Ocean*. Princeton University Press.

536 LIST OF FIGURES

537	Fig. 1. Fitted seasonal model ($\bar{\eta}$), deseasoned data (η') and original estimate from ECCOv4 (η)	28
538	Fig. 2. ARMA(n, m) demonstration	29
539	Fig. 3. Global η StD	30
540	Fig. 4. Percentage of variance in $\bar{\eta}$	31
541	Fig. 5. The $\bar{\eta}$ variance contribution over entire time interval (cm)	32
542	Fig. 6. Percentage of variance in η'	33
543	Fig. 7. Chosen order of ARMA(n, m) for the η'^\dagger and deseasoned η'	34
544	Fig. 8. Test for normality for η'^\dagger and η'	35
545	Fig. 9. Expected prediction error growth examples for η'^\dagger	36
546	Fig. 10. ARMA(n, m) expected prediction performance of η'^\dagger	37
547	Fig. 11. Test for normality for the annually averaged $\eta'^d ag$	38
548	Fig. 12. ARMA(n, m) expected prediction performance of annually averaged η'^\dagger	39
549	Fig. 13. Order of ARMA(n, m) using deseasoned η'^\dagger and η' , both annually averaged	40
550	Fig. 14. The chosen order of ARMA(n, m) using the BIC for $\eta'^d ag$	41
551	Fig. 15. ARMA(n, m) expected prediction performance of η' with the BIC	42
552	Fig. 16. ARMA(n, m) expected prediction performance difference of AIC and BIC	43

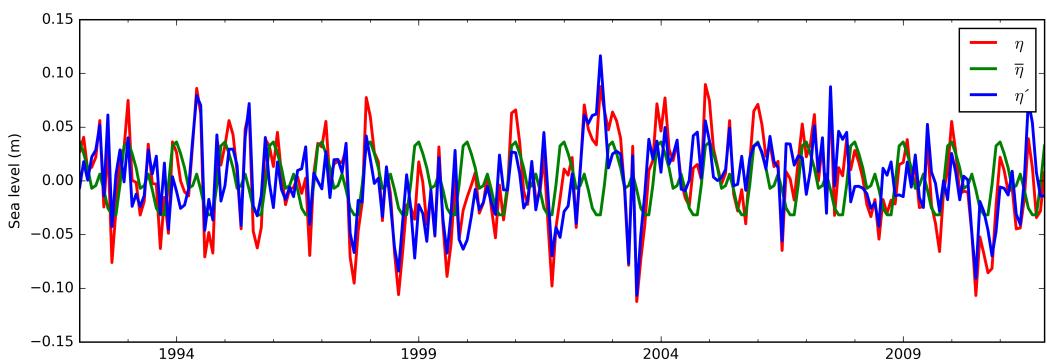


FIG. 1: Example of the process of removing the seasonal signal at 67°S , 149°W . The mean is removed, and the green line illustrates the fitted seasonal model ($\bar{\eta}$), the blue line are the deseasoned data (η'), while the red line shows the original estimate from ECCOv4 (η).

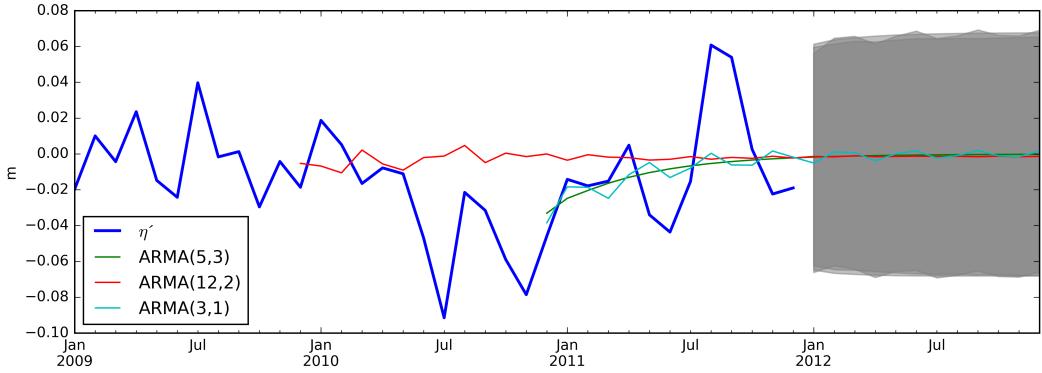


FIG. 2: Demonstration of the ARMA(n, m) models with detrended time-series of η' from 67°S , 149°W where the seasonal signal is removed. Three sample forecasts with different combinations of ARMA(n, m) orders are shown, with n and m respectively being (5,3) for the green line, (12,2) for the red line and (3,1) for the cyan line. The assumed noise is kept the same, and the gray envelopes show the standard deviations. Note that the projections approach the zero mean, but asymptote as different rates.

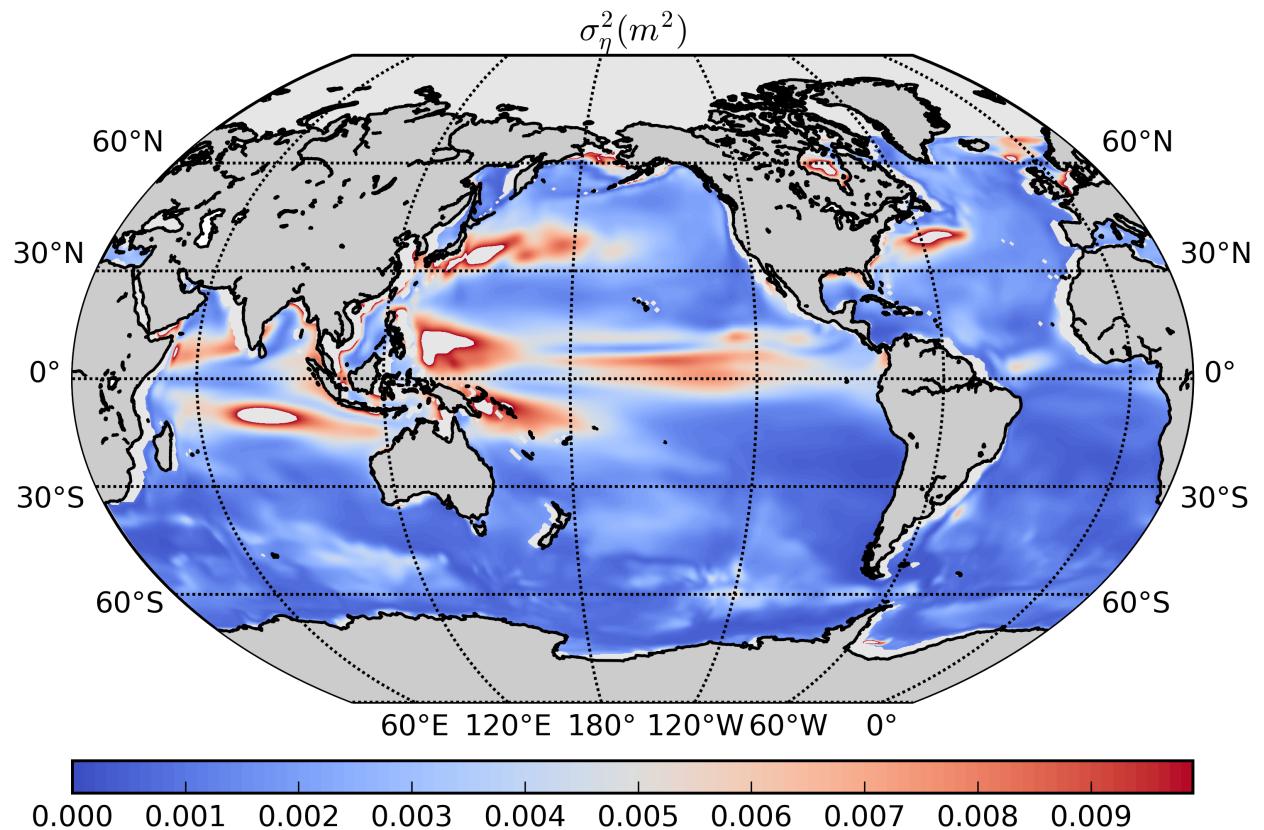


FIG. 3: Variance of η (m^2) from ECCOv4 from 1992-2011. Note the high variance associated with equatorial, western boundary currents and monsoonal regions.

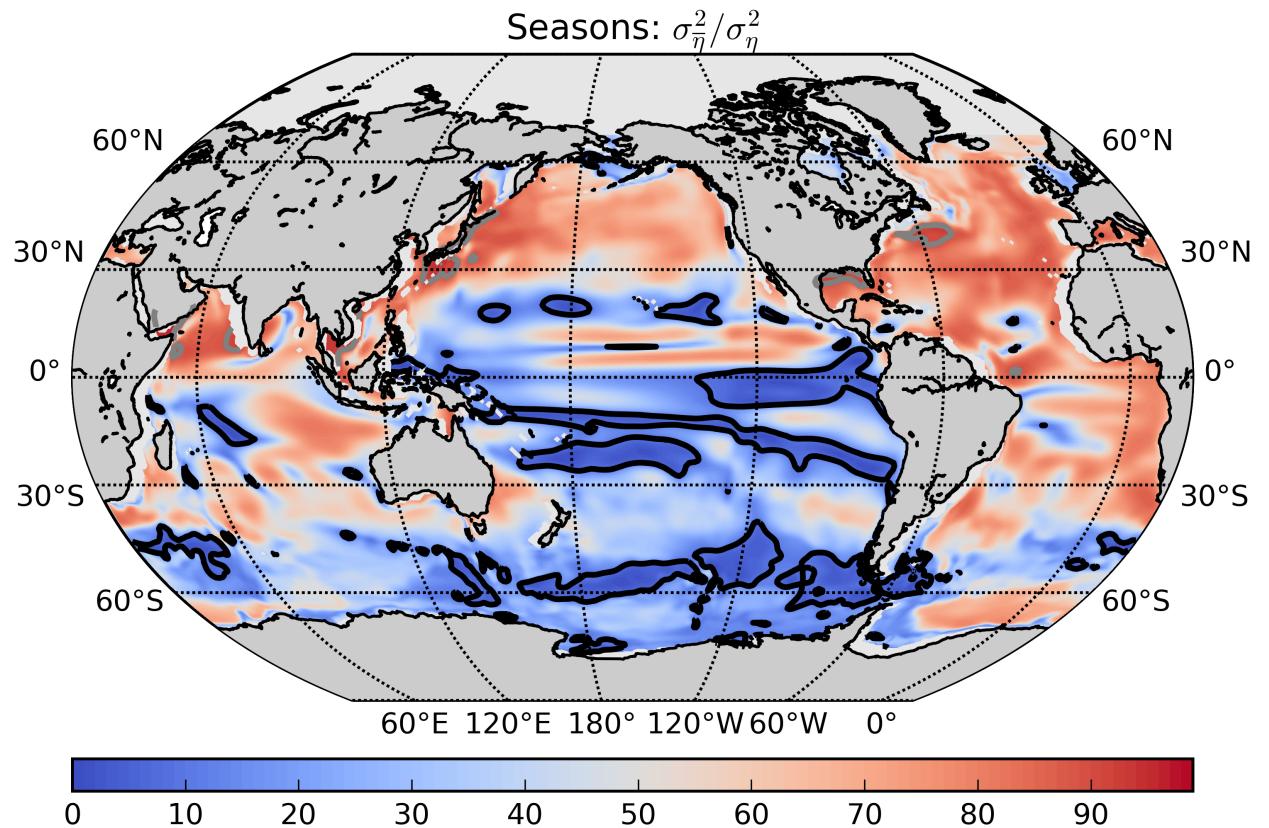


FIG. 4: Percentage of variance housed in seasonal η from ECCOv4 from 1992-2011 ($\sigma_{\bar{\eta}}^2 / \sigma_{\eta}^2$). Contours correspond to 5% (black) and 95% (grey) of values.

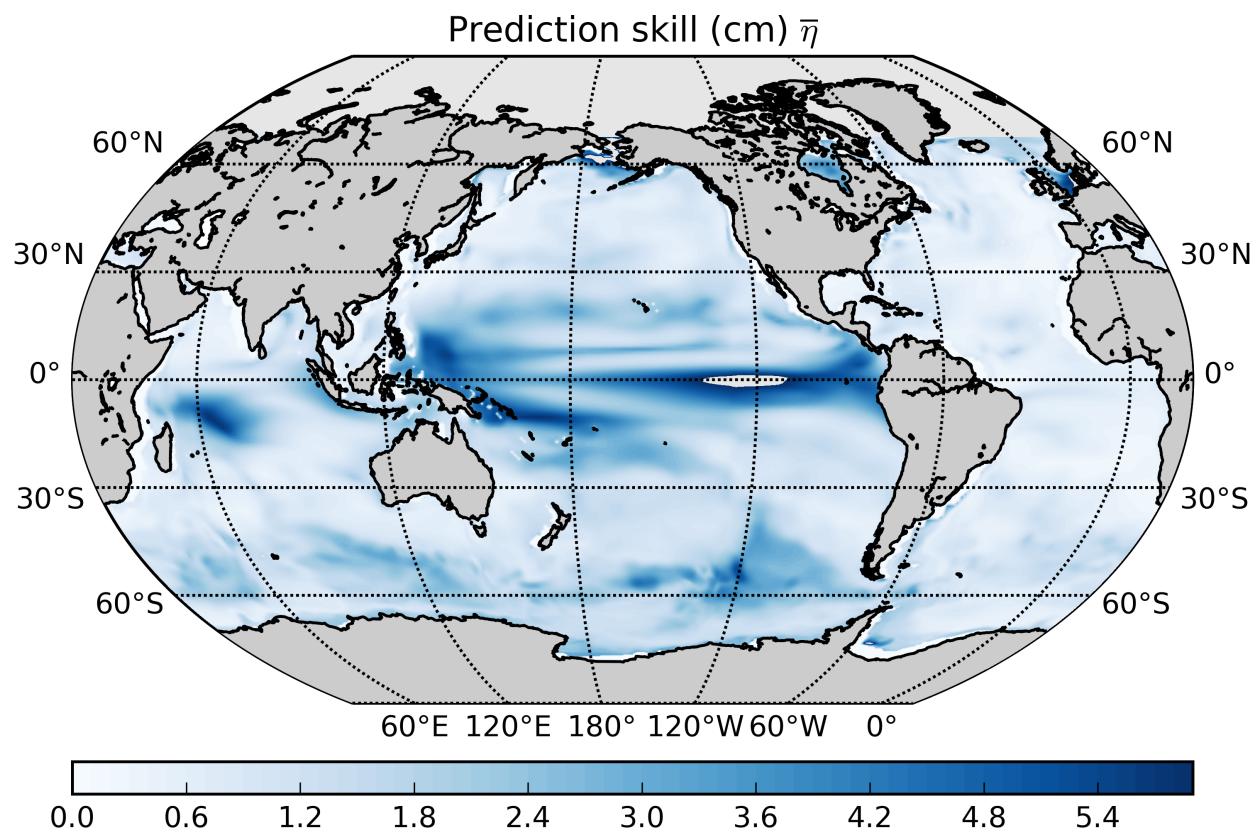


FIG. 5: The seasonal component, $\bar{\eta}$, is perfectly predictable, and is expressed as its contribution to the variance over the entire time interval in cm.

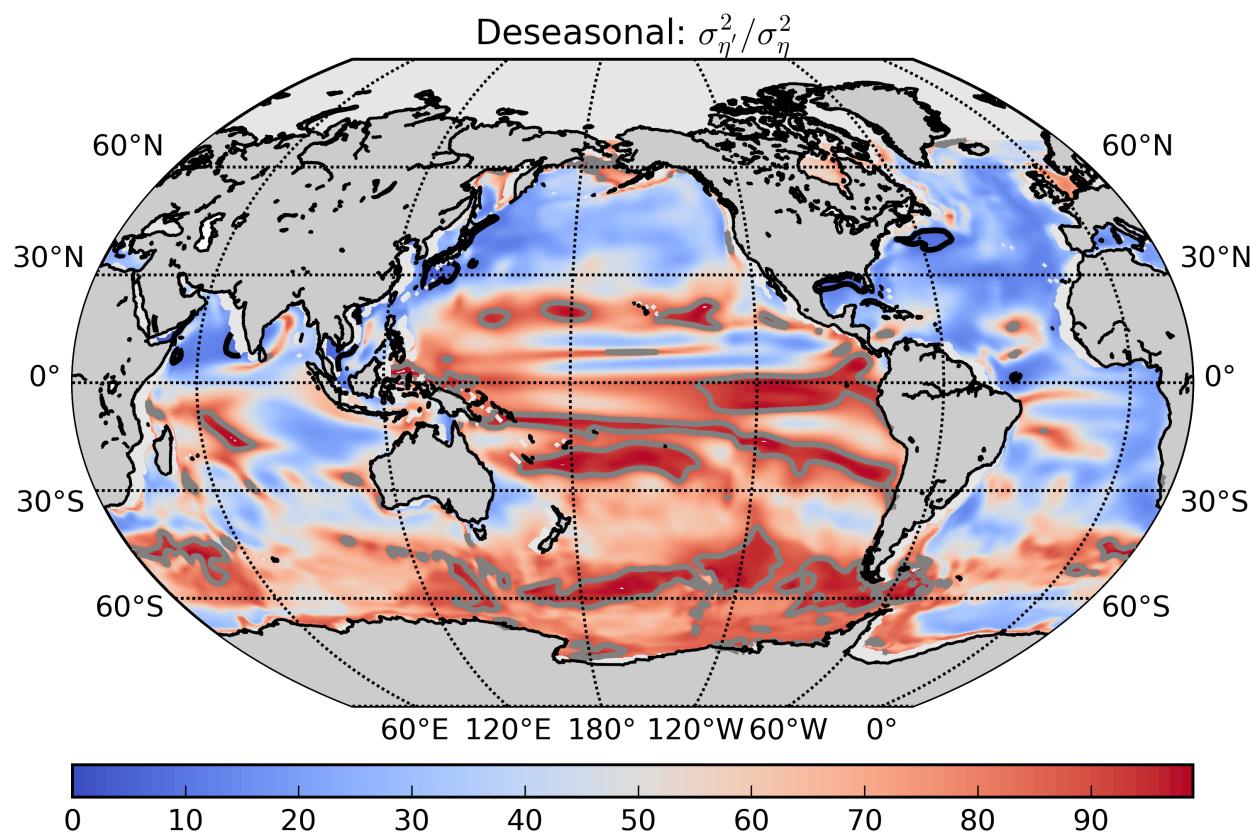


FIG. 6: Percentage of variance contained in η' from ECCOv4 from 1992-2011 ($\sigma_{\eta'}^2 / \sigma_{\eta}^2$). Contours correspond to 5% (black) and 95% (grey) of values.

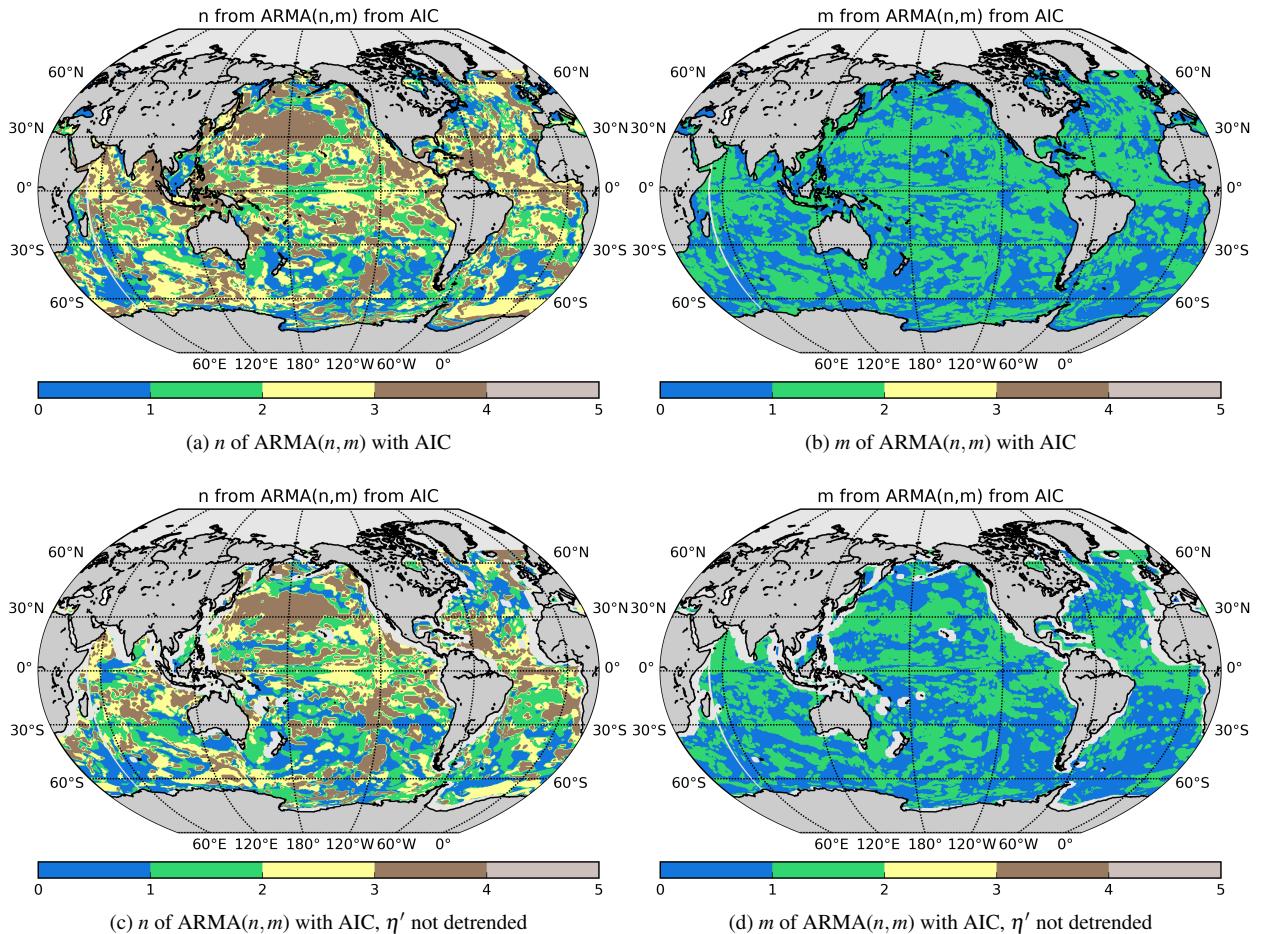


FIG. 7: The chosen order of ARMA(n,m) using the AIC for the η'^{\dagger} 1992–2011 (figures 7a and 7b showing the n and m of the ARMA(n,m), respectively), and for deseasoned η' data with the linear trend not removed (figures 7c and 7d showing the n and m of the ARMA(n,m), respectively).

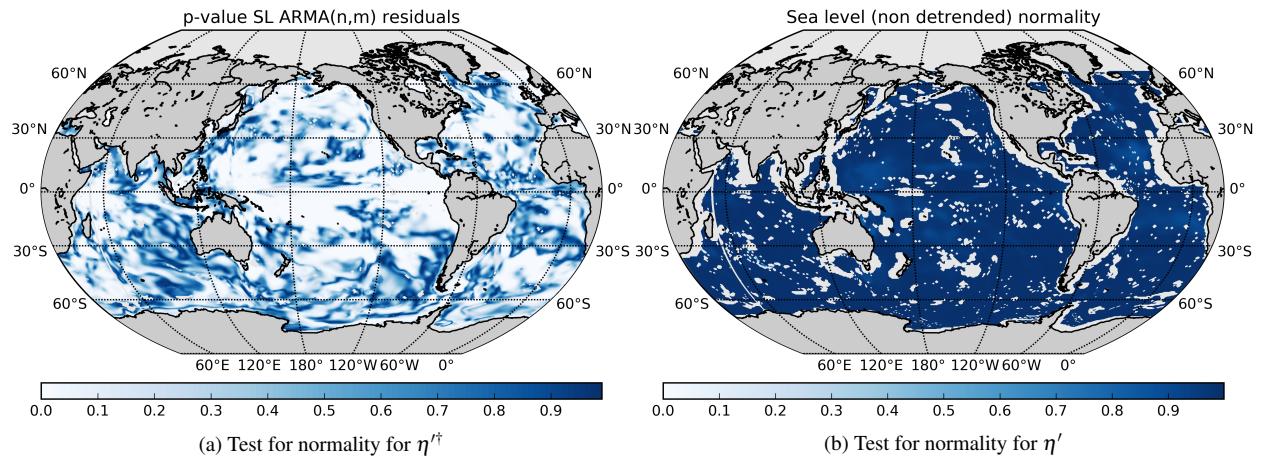


FIG. 8: Shapiro-Wilk test for normality for η'^\dagger (8a) and η' (8b). Darker blue indicating increasing confidence in accepting the null-hypothesis that the data is from a normally distributed population. White areas indicate areas where the AR coefficient roots were outside the unit circle, or where the MA coefficients were non-invertible.

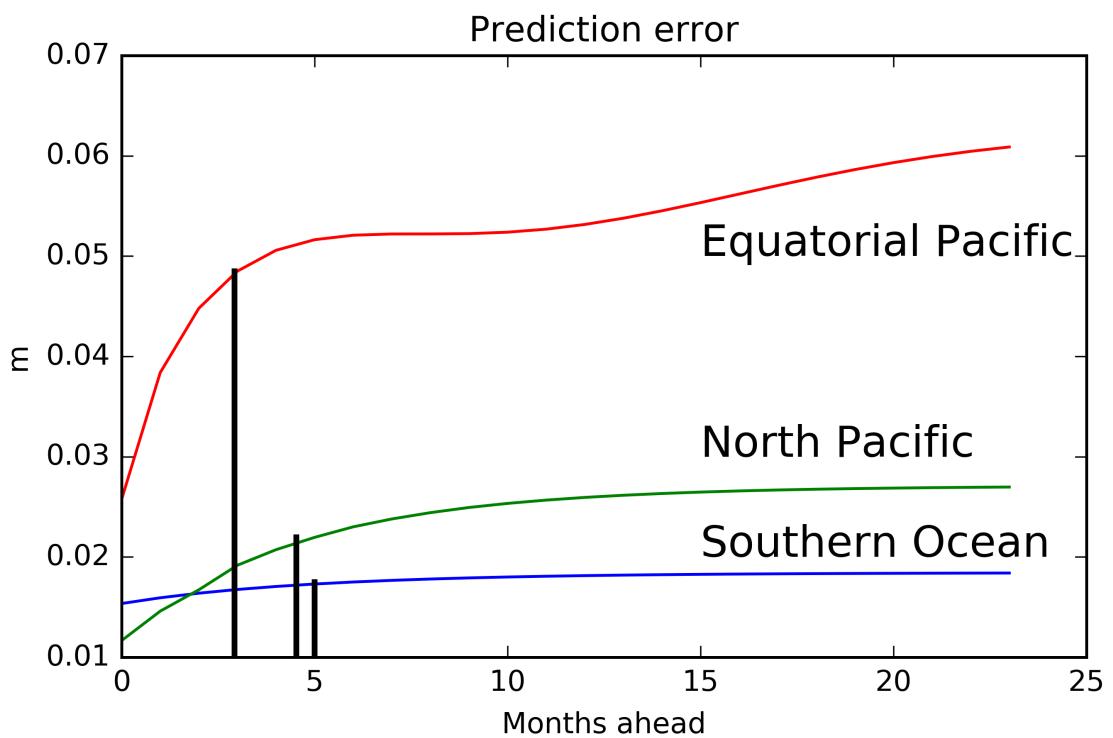


FIG. 9: The prediction error as defined in equation 5 for η'^{\dagger} in three locations in the Southern Ocean (blue line), North Pacific (green line) and the Equatorial Pacific (red line). The rate of error growth varies regionally, showing the e-folding timescale (black bars).

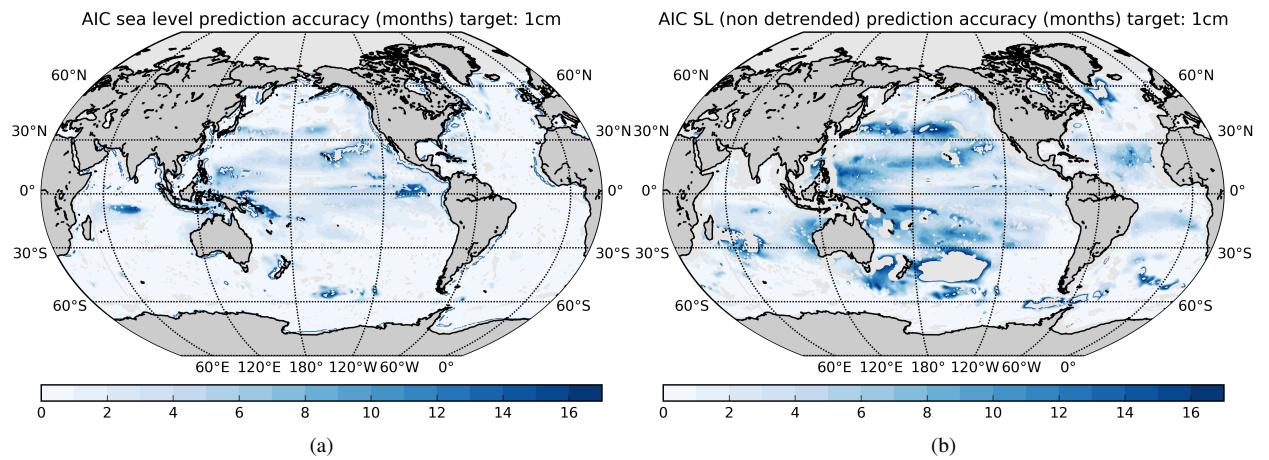


FIG. 10: The ARMA(n, m) expected prediction performance of η'^\dagger (10a) and η' (10b) phrased in terms of the time it takes the expected prediction performance to be less than 1 cm (months). Areas outside the colorscale indicate predictability longer than 17 months.

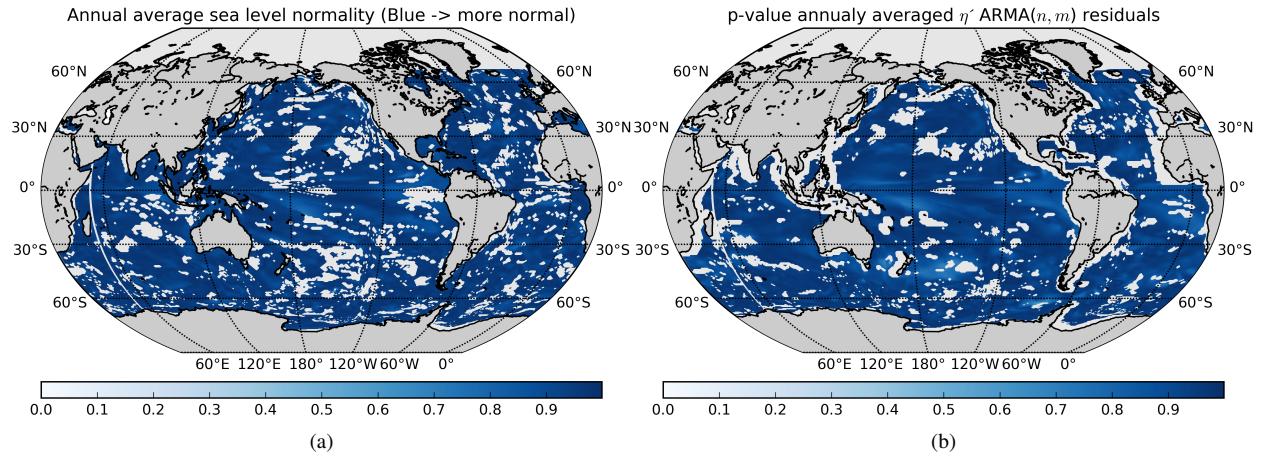


FIG. 11: The Shapiro-Wilk test for normality for the annually averaged η'^\dagger (11a) and η' (11b) data where the linear trend is not removed. Darker blue indicating increasing confidence in accepting the null-hypothesis that the data is from a normally distributed population. White areas indicate areas where the AR coefficient roots were outside the unit circle, or where the MA coefficients were non-invertible consistent with non-stationarity, likely due to noise.

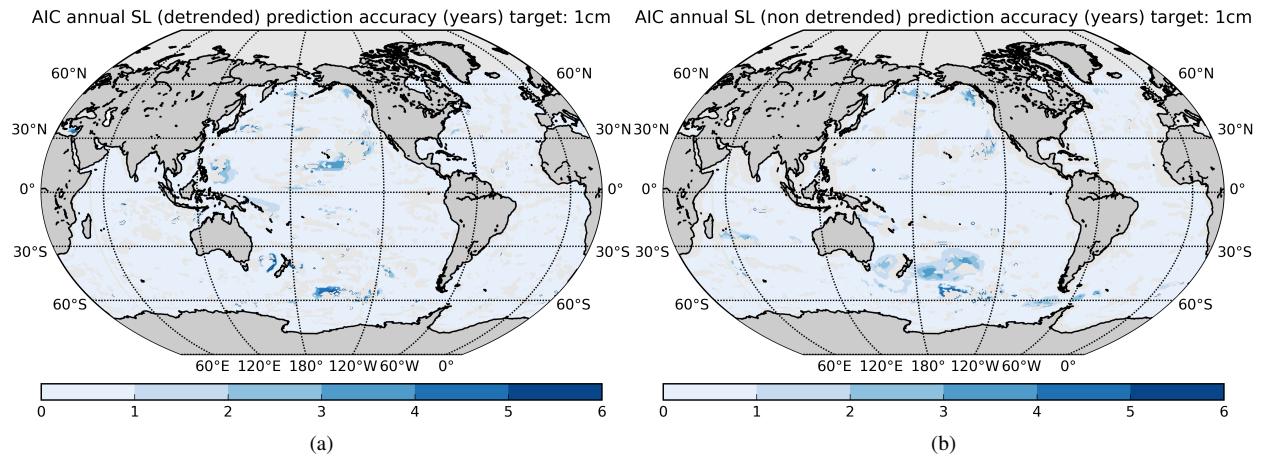


FIG. 12: The prediction performance based on the $\text{ARMA}(n, m)$ expected prediction error for annually averaged and η'^{\dagger} (12a) and η' (12b). Performance is phrased in terms of the time it takes the prediction performance to be less than 1 cm (years). Areas outside the colorscale indicate non-stationary ARMA coefficients.

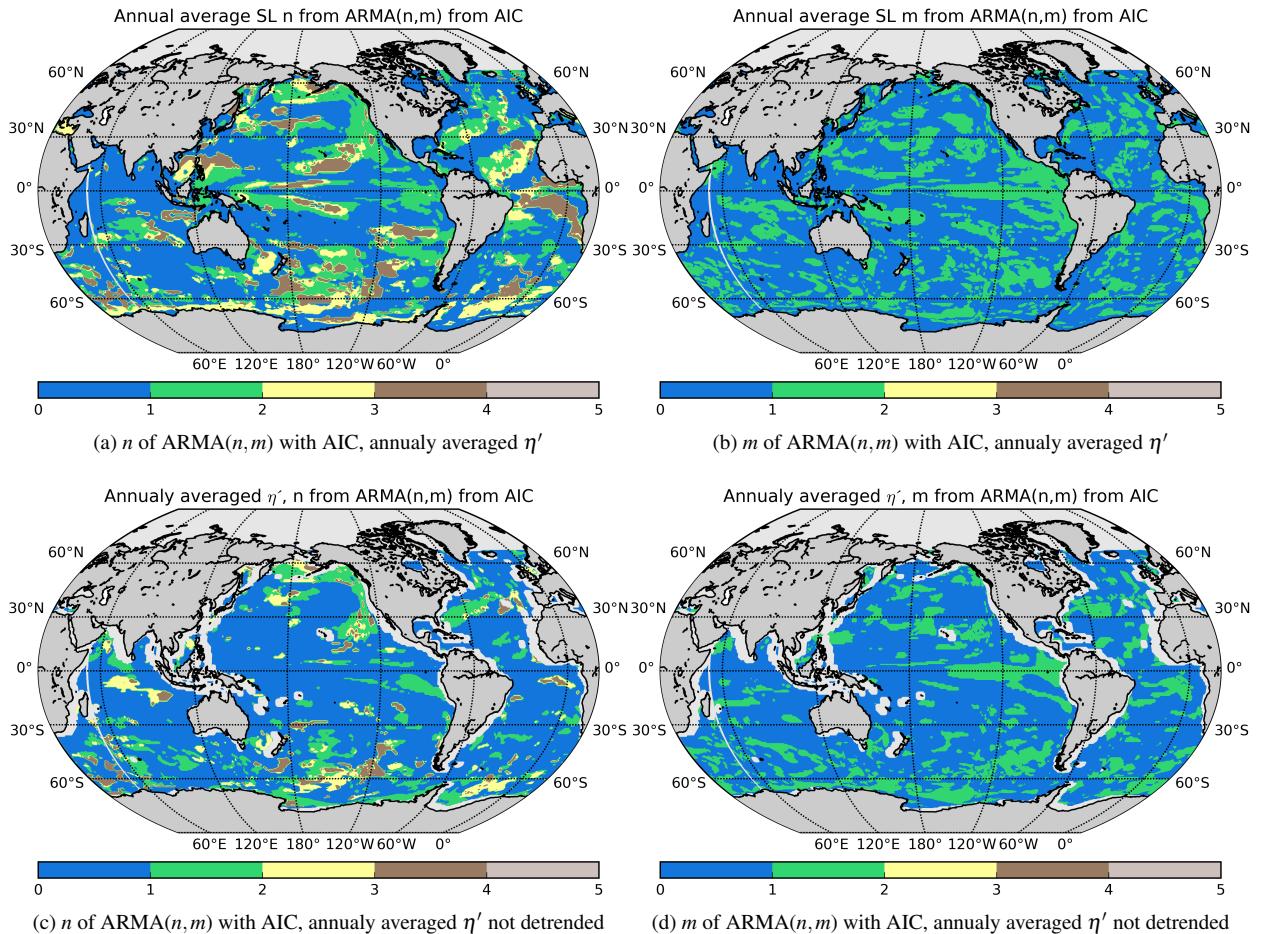


FIG. 13: Chosen order of ARMA(n, m) using the AIC of deseasoned η'^{\dagger} annually averaged with the linear trend removed (figures 13a and 13b showing the n and m of the ARMA(n, m), respectively), and η' annually averaged with the linear trend not removed (Figures 13c and 13d showing the n and m of the ARMA(n, m), respectively).

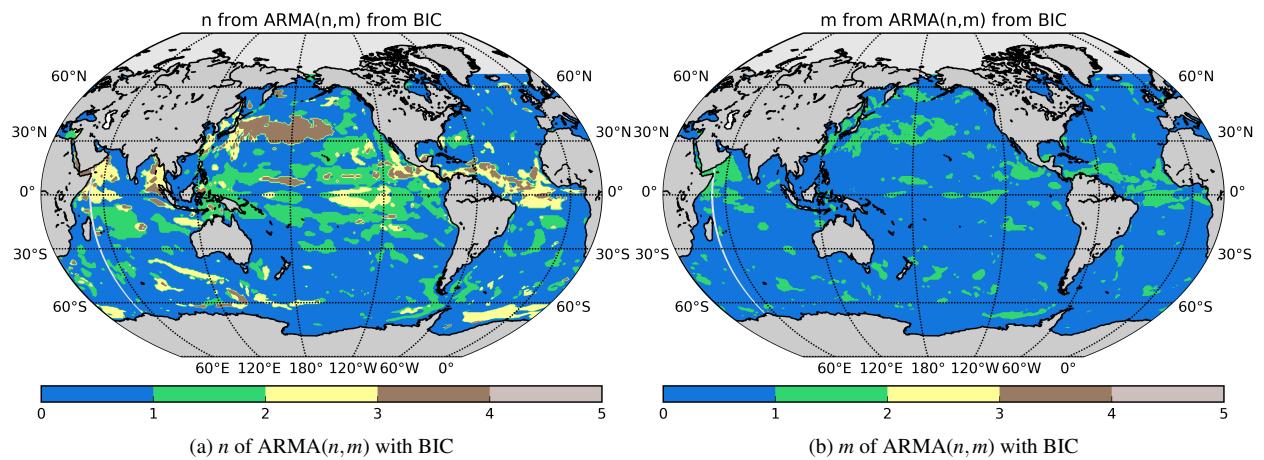


FIG. 14: The chosen order of ARMA(n, m) using the BIC (figures 14a and 14b showing the n and m of the ARMA(n, m), respectively). We use a four-point smoother for the η^{\dagger} 1992–2011. Note the difference to Figure 7 where higher orders are seen.

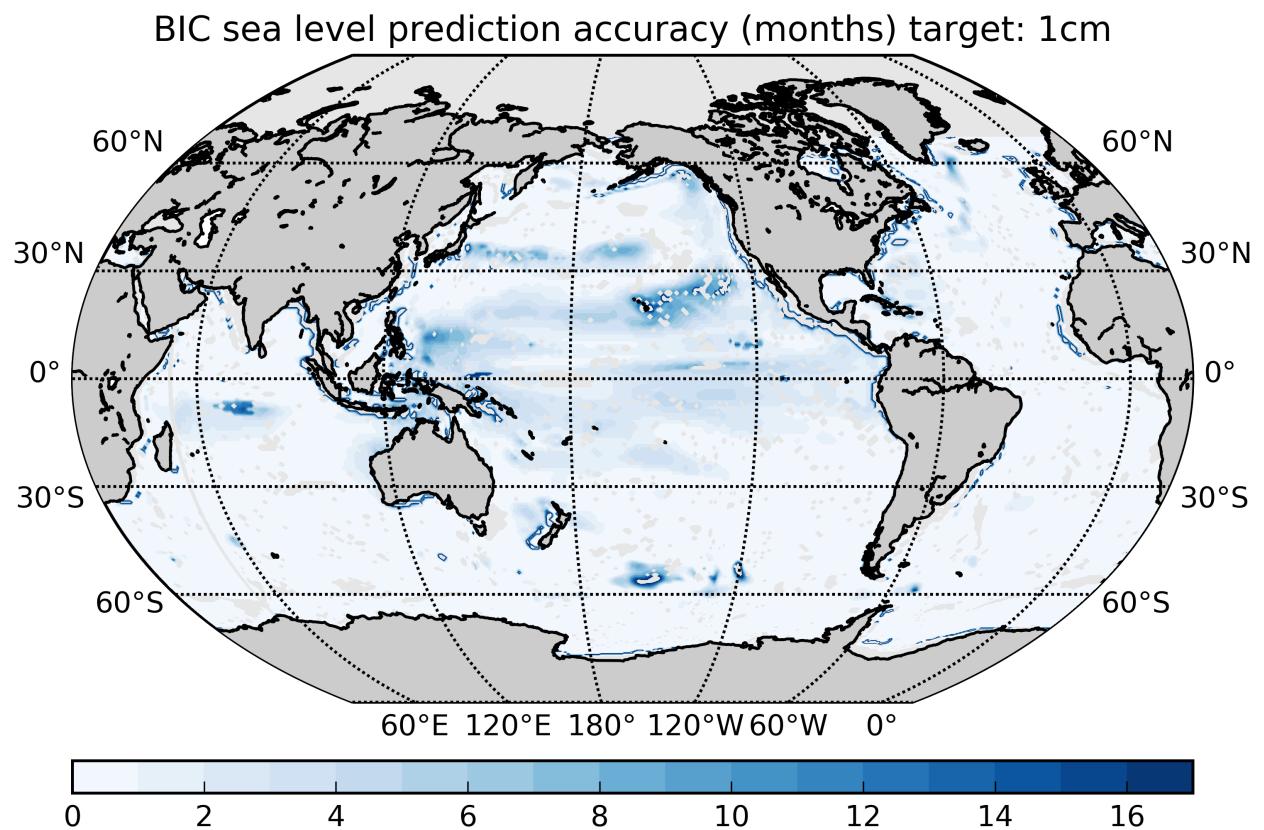


FIG. 15: The ARMA(n, m) expected prediction performance with BIC estimating the order of ARMA(n, m) models of η^{\dagger} phrased in terms of the time it takes the prediction performance to be less than 1 cm (months).

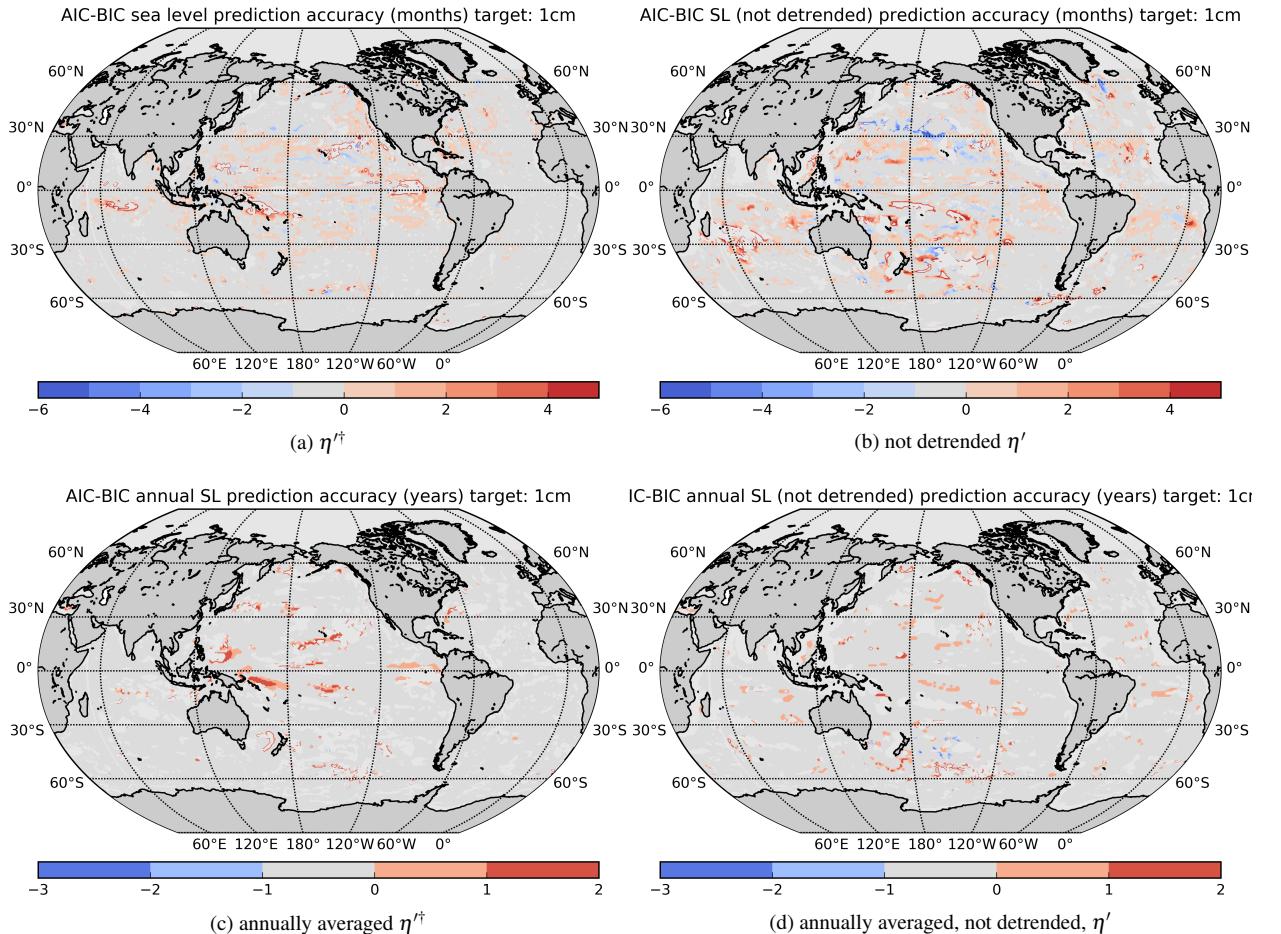


FIG. 16: The prediction performance difference with AIC-BIC using monthly averaged η'^\dagger (16a, in months), monthly averaged η' (16b, in months), annually averaged η'^\dagger (16c, in years) and annually averaged η' (16d, in years). Note that the AIC offers better predictability (red predominates) in most areas, but some prefer the BIC (blue).