

<sup>1</sup> **Linear predictability: A sea surface height case study**

<sup>2</sup> Maike Sonnewald\* and Carl Wunsch<sup>†</sup>

<sup>3</sup> *Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139*

<sup>4</sup> Patrick Heimbach

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

<sup>6</sup> *Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA*

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

<sup>9</sup> E-mail: maike\_s@mit.edu

<sup>10</sup> <sup>†</sup>Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02139, USA and  
<sup>11</sup> 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  $\eta$ ) 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 SSH. These factors include the wind field, atmospheric pressure, tides, ice-melt,  
34 river runoff, heat and freshwater exchange, and the shifting ocean circulation itself (Parker 1991;  
35 Church *et al.* 2013). The diverse physics spans a large range of timescales for oceanic response  
36 (e.g., Wunsch (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 many of the major components  
41 determining future values of  $\eta$  can be predicted from a knowledge of the present and past states  
42 of the ocean. The expected prediction error (PE) growth of  $\eta$  is not well established. Attempting  
43 to estimate the PE via ensembles of climate model simulations reveals large ensemble spread.  
44 (e.g., the Intergovernmental Panel on Climate Change, IPCC) (Church *et al.* 2013; Stainforth *et*  
45 *al.* 2005; Palmer 2012).

46

47 The goal here is to assess quantitatively the extent to which  $\eta$  variability is predictable using  
48 linear methods, describing both the deterministic (seasonal changes) as well as the underlying  
49 continuum treated as a wide-sense stationary linear process.<sup>1</sup> As discussed e.g., by Wunsch  
50 (2013), such an approach provides a baseline against which predictions made with consider-

---

<sup>1</sup>“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 ably more complex methods (non-linear, non-stationary, extended to spatial structure) can be  
52 compared. The general case involves much more complex computations, and raises the purely  
53 practical issue of whether the linear, univariate, stationary approach is adequate for SSH and for  
54 how long?

55

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

64

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

69

70 Interpretation of statistics from short records is difficult (see e.g., Wunsch (1999); Percival *et al.*  
71 (2001); Ocaña *et al.* (2016)). The methods that underlie much of what is presented here rely on  
72 the assumptions that  $\eta$  changes from the superposition of deterministic seasonal components and  
73 a wide-sense stationary stochastic process. Of most relevance for the latter are general red-noise  
74 processes and the extreme of white noise, which is by definition linearly unpredictable. Detection

75 of true non-stationarity is not possible with the short records at hand. Similarly, an infinite number  
76 of generalizations to non-linear representations are possible, but unless the linear assumption can  
77 be excluded, it remains an important reference point.

78

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

87

88 The methods are detailed in Section 2 and the results are presented in Section 3, where the sea-  
89 sonal and non-seasonal contributions to the variance of  $\eta$  are presented. As defined in this paper,  
90 the seasonal component is perfectly predictable and the non-seasonal portion involves stochastic  
91 forecasting. A set of four restimates is presented: using monthly or annual means of  $\eta$ ; with ap-  
92 parent linear trends included as part of the background rednoise; and with the trends removed in  
93 both cases. Section 4 presents the discussion and conclusion.

94 **2. Numerical and ARMA models**

95 Predictability of SSH is studied using the ECCOv4 global bi-decadal state estimate, as described  
96 by Wunsch and Heimbach (2013); Forget *et al.* (2015) and others (see also ECCO Consortium  
97 (2017a,b)). The state estimate is global  $1^\circ$  with tropical mesh refinement. A least-squares with

<sup>98</sup> Lagrange multipliers approach is used to obtain the state estimate. The result is an adjusted, yet  
<sup>99</sup> *free-running* version of the MIT General Circulation Model (MITgcm, Adcroft *et al.* (2004)). In  
<sup>100</sup> contrast to most “reanalysis” products, the ECCO oceanic state satisfies basic conservation laws  
<sup>101</sup> for enthalpy, salt, volume, and momentum remaining largely within error estimates of a diverse  
<sup>102</sup> set of global data (Wunsch and Heimbach 2007, 2013; Stammer *et al.* 2016). Regions without  
<sup>103</sup> data are filled in a dynamically consistent way using the dynamics, avoiding the use of untested  
<sup>104</sup> statistical hypotheses e.g., Reynolds *et al.* (2013).

<sup>105</sup> At each point of latitude and longitude  $(\theta, \lambda)$  of the state estimate, the temporal mean (1992-  
<sup>106</sup> 2011) is removed and  $\eta(\theta, \lambda)$  is defined as  $\eta = \eta'(\theta, \lambda) + \bar{\eta}(\theta, \lambda)$ , where  $\bar{\eta}(\theta, \lambda)$  denotes the  
<sup>107</sup> seasonal and  $\eta'(\theta, \lambda)$  the non-seasonal  $\eta(\theta, \lambda)$ . Throughout this study, each  $\theta, \lambda$  coordinate is  
<sup>108</sup> used and for simplicity the spatial indicies are dropped hereafter. At this stage, any temporal trend  
<sup>109</sup> is being included as part of  $\eta'$ .

<sup>110</sup>

<sup>111</sup> The seasonal component here includes its first two harmonics as illustrated in Figure 1. Variances  
<sup>112</sup> of the seasonal and non-seasonal components are additive:

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

<sup>113</sup> To the extent that  $\sigma_{\eta}^2 \gg \sigma_{\eta'}^2$ , useful prediction is purely deterministic. When the seasonal  
<sup>114</sup> variability is not dominant, the predictability of the non-seasonal process has to be examined.  
<sup>115</sup> Deterministic prediction of sinusoidal components is straightforward.

<sup>116</sup>

<sup>117</sup> Linear predictability of wide-sense stationary stochastic processes, not distinguishable from  
<sup>118</sup> Gaussian, is well-understood with a very large literature including standard textbooks (e.g., Box  
<sup>119</sup> and Jenkins (1970)). Here the formalism is discussed only insofar as it develops the notation to

120 be applied. Most linear methods are based on the autoregressive process of order  $n$  (AR( $n$ )), the  
 121 moving average process of order  $m$ ,(MA( $m$ )), or the mixed autoregressive moving average of  
 122 order  $n,m$  (ARMA( $n,m$ )).

123

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

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

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

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

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

128 where  $a_i$  and  $b_i$  are regression coefficients,  $e(t)$  is near-Gaussian white noise with zero mean and  
 129 variance  $\sigma_e^2$ , and  $t$  is any time, past, present or future, measured in units producing an implied  
 130 time-step  $\Delta t = 1$ . Conversion of one form to another or to the mixed representation is discussed  
 131 in textbooks (e.g., Box and Jenkins (1970)). In the MA form, the white noise increments,  $e(t)$ ,  
 132 are determined for past values leading up to the present time,  $t$ .  $e(t)$  is known (estimated), but  
 133 no future values  $e(t + \Delta t)$ ,  $\Delta t > 0$  are known. In the AR form, past values  $\zeta(\tau)$  are assumed  
 134 to have been estimated, as is  $e(t)$ , but again no future values are available. In the presence  
 135 of noise, these representations can become unstable, being indistinguishable from the presence  
 136 of apparent non-stationarity. Tests for stability/non-stationarity are based upon the zeros and

<sup>137</sup> poles of complex polynomials formed from the various coefficients,  $a_i, b_j$  (Box and Jenkins 1970).

<sup>138</sup>

<sup>139</sup> The MA form gives the simplest representation of the growth of prediction error, from 5, al-  
<sup>140</sup> though the final growth rate is the same for all consistent ARMA forms. (All results here are  
<sup>141</sup> based on the conversion to the MA after determination of the more general ARMA.) Converting  
<sup>142</sup> the ARMA to the MA form using the Wold representation the  $\tau$ -ahead PE is:

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

<sup>143</sup> This quadratic error growth depends upon the values of  $b_p$ , whose sum can never exceed the time  
<sup>144</sup> series variance  $\langle \zeta^2 \rangle$ , for which the best prediction would be the time mean.

<sup>145</sup>

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

<sup>149</sup>

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

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

<sup>157</sup> where  $\mathcal{L}$  is the likelihood:

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

<sup>158</sup>  $\zeta_i$  is the observed, and  $\hat{\zeta}_i$  is the prediction, so  $(\zeta_i - \hat{\zeta}_i)^2$  are the prediction residuals. In the estimate,  
<sup>159</sup> the AIC value is minimized, which determines the smallest appropriate order to represent the  
<sup>160</sup> time-series. As discussed by Priestley (1981) and Yang (2005), the AIC can overestimate the  
<sup>161</sup> order; see Appendix A1 for more detail.

<sup>162</sup>

<sup>163</sup> In the following discussion of of the SSH time series, two cases are considered: one where only  
<sup>164</sup> a time mean has been removed ( $\eta$ ), and one where a best-fitting linear trend has been subtracted  
<sup>165</sup> as well. Separate analyses of ( $\eta^\dagger$ ) using both monthly and annual mean time series are considered.  
<sup>166</sup> Further removing a seasonal cycle  $\bar{\eta}$  leads to time series  $\eta'$  and  $\eta^{\dagger\dagger}$ , respectively. A significant  
<sup>167</sup> linear or quadratic trend can itself be used to make a prediction. By including trend structures in  
<sup>168</sup> the stochastic process, the predictability of the time series will be enhanced.

<sup>169</sup>

### <sup>170</sup> 3. Results

#### <sup>171</sup> a. Seasonal Variance

<sup>172</sup> Figure 2a shows the total variance  $\sigma_\eta^2$  of  $\eta$  from monthly means of 1992-2011. Western  
<sup>173</sup> boundary currents and their extensions are associated with higher variance, particularly in the  
<sup>174</sup> Northern Hemisphere. The tropics have a large SSH variance, particularly in the Pacific Ocean.

175 The Indian Ocean is dominated by monsoonal effects particularly in the Arabian Sea warm pool  
176 and complicated interactions of jet dynamic and the Indonesian throughflow seen particularly  
177 around 15°S (Schott *et al.* 2001). Elevated standard deviations occur in the eastern Indian  
178 Ocean, that are not seen in the Atlantic or Pacific Oceans. The Southern Ocean shows some  
179 excess variance, particularly in the Indian and Pacific sectors. Bathymetric features such as the  
180 Pacific-Antarctic ridge are associated with higher variance (Ponte and Piecuch 2014).

181

182 Figure 2b illustrates the percentage of the total variance included in the  $\bar{\eta}$  component. A  
183 striking, but well-known, seasonal hemispheric difference appears across the equator, where  
184 much of the  $\eta$  variance in the Northern Hemisphere is dominated by the seasonal component.  
185 Exceptions include a large zonal band in the Pacific and areas at a similar latitude in the North  
186 Atlantic. The Irminger and Labrador Seas also have large areas where the seasonal signal is less  
187 dominant, as is also true of the Bering Sea. The South Pacific Ocean has large areas where the  
188 seasonal signal is weak, but a signal extends westwards off the coast of Peru, likely associated  
189 with the upwelling there. A similar feature is seen in the East Australian Current. In the Indian  
190 Ocean, fluctuations in the Arabian Sea are almost entirely captured by the seasonal monsoonal  
191 component. South of the Equator, the eastern half of the Indian Ocean in the Southern Hemisphere  
192 is largely dominated by  $\bar{\eta}$ , but less clearly so in the West. In the Southern Ocean, a clear seasonal  
193 dominance appears in the Weddell Gyre region, as well as in the area of the standing meander off  
194 the Agulhas coast of South Africa.

195

196    *b. Seasonal Prediction*

197    The seasonal component,  $\bar{\eta}$ , is predictable as defined over the bi-decadal time interval covered  
198    by ECCOv4, and Figure 2d illustrates the associated contribution to the prediction if the StD of the  
199    seasonal component in cm over 1992-2011. As expected, where the seasonal variance is a large  
200    fraction of the total, good predictability is found, e.g associated with the seasonal component in  
201    the Mascarene basin area, as well as in the Southern Ocean in the Indian and Pacific Ocean sectors.

202

203    *c. Non-seasonal Variance*

204    Figure 2c shows the percentage of the total variance accounted for by the non-seasonal  
205    background process. Here a visually striking signal appears across the Pacific Ocean along  
206    the equator, probably associated with the El Niño Southern Oscillation (ENSO) climate mode.  
207    Further zonal bands appear to the north of the equator. A strong signal occurs in the Bering  
208    Strait region. In the Atlantic Ocean, the subtropical regions show active areas, and along the  
209    paths of the Labrador and southern tip of the East Greenland currents. In the South Atlantic, the  
210    non-seasonal component of variance accounts for a large portion of the variance in the Brazil  
211    current and Zapiola region of the Argentine Basin, as well as in zonal bands. In the Indian Ocean,  
212    the non-seasonal component of variance also accounts for a large portion of the total variance to  
213    the west in the Mascarene basin region, as well as along the western Indonesian coast associated  
214    with the propagation of the throughflow. The Southern Ocean produces a very large signal  
215    associated with regions of deep mixed layers and possible mode water formation, and in areas  
216    where the ACC is directed southwards. Predictability associated with this non-seasonal variance  
217    is addressed in the remainder of this paper.

218

219 *d. Predictability After Trend Removal*

220 A linear trend is now removed from the SSH values, meaning that a possibly perfectly  
221 predictable component is eliminated. The  $\eta'^\dagger$  is fit to an ARMA( $n, m$ ) process of the ECCOv4  
222 state estimate from 1992-2011. A four point smoother is applied to  $\eta'^\dagger$  to reduce noise (two points  
223 in latitude and two in longitude), equivalent to moving from the tracer ( $t$ -point) to the vorticity  
224 point ( $f$ -point) in the Arakawa C-grid. Using the smoother tends to make the data adhere more  
225 closely to a normal distribution, but can exaggerate the spatial covariance of isolated outliers. The  
226 performance of the different choices is given by the rate of the PE growth over time. Figure 3a  
227 show the order  $n$  of the ARMA( $n, m$ ) chosen with the AIC. The order chosen shows how many  
228 coefficients are used to optimally represent  $\eta'^\dagger$  and the associated prediction error.

229

230 The simplest linear theory assumes that the underlying values are Gaussian or close to it, an  
231 assumption tested in the ECCO estimated SSH in Figure 4a using the Shapiro-Wilk test for  
232 normality (Shapiro and Wilk 1965). Large areas associated with features such as the ENSO signal  
233 appear to deviate from normality (i.e., p values close to 0). This result has implications for the  
234 predictability since these departures are important when interpreting the PE.

235

236 Figure 5 shows the error growth asymptoting to its upper-bound: the full variance of the  
237 background residual time series. Large differences as a function of region appear in the PE, as  
238 well as in their asymptotic rate of growth. The expected error  $e$ -folding structure is shown to  
239 illustrate the rate of predictability decay, independent of magnitude.

240

241 The ARMA( $n, m$ ) expected PE growth associated with the  $\eta'^\dagger$  is illustrated in Figure 6a. The  
242 PE growth is expressed in terms of the time taken (months) before a target is exceeded (1cm).  
243 Large areas of the ocean show limited performance based on the PE growth over time, but certain  
244 (mostly isolated) areas have good performance, exceeding a year. In interpreting these Figures,  
245 it is prudent to keep in mind the assumption of normality(Figure 4a). High predictability is seen  
246 in the North Pacific Ocean. Although the physics here are beyond the intended present scope  
247 of this paper, these areas are associated with the Kuroshio crossing the Pacific and wave-eddy  
248 interactions stretching from Hawaii to the coast of California in a banana-shape. Areas in  
249 the Equatorial Pacific also show better performance of PE growth over time, but are clearly  
250 non-normal. Features associated with the Pacific-Antarctic Ridge in the Southern Ocean also  
251 produce better performance in terms of the ARMA( $n, m$ ) expected PE growth over time, along  
252 with some areas in the Irminger 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  
259 *et al.* 2014; Ocaña *et al.* 2016). To assess the impact of including the trend, analysis of the  
260 background process is repeated. Figures 3c and 3d illustrates the ARMA( $n, m$ ) order chosen  
261 by the AIC. This result is similar to Figures 3a and3b, but sometimes smaller values of  $n, m$   
262 result—as is physically plausible with a trend.

263

264 The normality of the stochastic background process with the trend is re-tested in Figure 4b, illus-  
265 trating that most of the ocean remains indistinguishable from a normal distribution in  $\eta'$ . Figure  
266 6b shows the prediction performance based on the ARMA( $n, m$ ), phrasing the result in terms of  
267 the number of months it takes for the PE to cross the accuracy target of 1 cm. Retaining the trend  
268 adds predictability, with large areas taking over 12 months before exceeding the 1 cm threshold.  
269 Areas where the trend is important are in bands in the subtropics in both the Atlantic and Pacific,  
270 as well as large areas off the coast of Greenland, the Drake passage, the Kuroshio path across the  
271 Pacific, the southern Indian Ocean and a remarkably large region in the South Pacific poleward of  
272 30°S. Largely, Figure 6b is seen to amplify Figure 6a, but with notable exceptions. This could be  
273 associated with the spreading of a thermosteric signal, particularly the the subtropical South Pa-  
274 cific associated with PDO dynamics. The subtropical Atlantic stands out as another region where  
275 the trend is key, along with the South Pacific. The mechanisms, particularly in terms of linear  
276 dynamics, are not clear. For example, regions associated with bathymetric features like the Pacific  
277 Antarctic Ridge do not stand out as intuitive regions of heat storage.

278 *f. Predictability with Annual Averages*

279 Interannual and monthly physics are distinct. Assessing the annually averaged  $\eta'^\dagger$  and  $\eta'$  from  
280 ECCOv4 separately, the assumptions of estimating the covariance and of a Gaussian distribution  
281 are likely inaccurate, owing to the short, 20-value record.

282

283 Figure 9 shows the chosen ARMA( $n, m$ ) order. A four-point smoother is again used for reduced  
284 noise on the ECCOv4  $\eta'^\dagger$  1992-2011 where the trend is removed and the data are annually  
285 averaged. The  $n$  and  $m$  of the ARMA( $n, m$ ) now reflect the annually averaged data, rather than  
286 the monthly. The  $n$  of the ARMA( $n, m$ ) is generally lower for the annually averaged  $\eta'^\dagger$  than for

287 monthly  $\eta'^\dagger$ . However, large regions show very different patterns than from the monthly data. Ex-  
288 amples include the Arabian Sea region and the North Pacific. However, the areas where one order  
289 dominates are generally larger than for monthly  $\eta'^\dagger$ . Regions that have higher orders for yearly  
290 averaged data than for monthly averages are areas such as the Pacific sector of the Southern Ocean.

291

292 Figure 7a assesses the extent to which the 20-year time series can be viewed as coming from  
293 a normally distributed population when using annual averages. Most of the ocean passes this  
294 test for annually averaged  $\eta'^\dagger$ , but with large areas that appear noisy/non-stationary, presumably  
295 owing primarily to the presence of noise.

296

297 The associated prediction performance based on the ARMA( $n, m$ ) expected PE growth over  
298 time is shown in Figure 8a. In most areas the time to reach the criterion of 1 cm expected  
299 error is one years, but areas exceeding 3 are seen. These regions are clustered in the Pacific,  
300 with a patch in the western equatorial Pacific, a band stretching from the central North Pa-  
301 cific to the coast of the USA (Hawaii to California), as well as isolated small patches elsewhere.  
302 A larger patch of good PE performance over time exists in the Pacific sector of the Southern Ocean.

303

304 As expected, the role of the linear trend is also important in the annually averaged  $\eta'$ . Figure  
305 9a-c illustrates the associated orders of the ARMA( $n, m$ ). As with the monthly data, the orders are  
306 generally lower. Higher orders exist in a band stretching from the central Pacific northwards to  
307 the USA (Hawaii to California) as well as a feature associated with the Antarctic-Pacific ridge.

308

309 Figure 7b using annually averaged  $\eta'$  suggests that most of the ocean passes the test of  
310 normality, but again with large areas of failure. The associated prediction performance based on

311 the ARMA( $n, m$ ) expected PE growth over time is shown in Figure 8b. This result is similar to  
312 that in Figure 8a, but most of the longer-term PE performance over time is found in the South  
313 Pacific. A band stretching from the central Pacific northwards to the USA still exists, but large  
314 areas do not pass the test of a stable or wide-sense stationary ARMA. Patches of longer term PE  
315 performance over time are also seen in the Indian Ocean and Drake Passage, with isolated regions  
316 elsewhere.

317

#### 318 **4. Discussion and Conclusions**

319 In this paper, linear univariate predictability of SSH,  $\eta$ , is discussed, and benchmarks for more  
320 elaborate prediction methods are presented. In general, more complex models and prediction  
321 methods (non-linear, non-normal) schemes would need to exceed this PE performance over time  
322 to be proven worthwhile. Prediction performance is presented in terms of the time it takes the  
323 expected prediction error (PE) to grow beyond 1 cm. More complex models should necessarily do  
324 better and their use may well be justified, particularly in specific physically identifiable regions.  
325 This approach is supported by work such as Goddard *et al.* (2015), where certain events in  $\eta'$  have  
326 been attributed to factors such as changes in the Atlantic meridional overturning circulation and  
327 the North Atlantic Oscillation. Existing spreads in ensemble studies, such as the CMIP5 models,  
328 would suggest that many difficulties remain (Church *et al.* 2013).

329

330 In the ECCOv4 state estimate, the seasonal cycle,  $\bar{\eta}$  seen in Figure 2b, accounts for more than  
331 80% of the variability over 20+ years in large parts of the Atlantic and Northern Hemisphere  
332 Pacific, as well as the Weddell Gyre area in the Southern Ocean. In these regions, the seasonal

333 component is likely sufficient for estimating the variation in  $\eta$  for at least a few decades.

334

335 The percentage of 20-year variance in the stochastic non-seasonal component,  $\eta'$  seen in Figure  
336 2c, is important over large areas of the ocean, particularly in the Southern Hemisphere.  $\eta'$  is  
337 treated as a weakly stationary univariate random process, with the assumption of being normally  
338 distributed. Areas where  $\eta'$  is particularly important for predictions, as seen in Figure 6b, lie in  
339 the Southern Ocean, and in extensive regions of the Pacific as well as in the Indian Oceans. When  
340 a linear trend is removed as in Figure 6a, regions where the ARMA prediction still performs well  
341 are mainly clustered in the Pacific, particularly in a band extending northeast from Hawaii, as well  
342 as in isolated small areas elsewhere. A patch of higher predictability exists in the Pacific sector  
343 of the Southern Ocean. However, the ARMA prediction, as expected, tends to perform less well  
344 when the trend is removed.

345

346 An important next step is to distinguish the physical mechanisms, whether atmospheric or  
347 oceanic,in regions where the ARMA prediction procedure does well and those where it works  
348 poorly. Stationary linear prediction methods can be of limited utility for a variety of reasons.  
349 These include dominance by unpredictable white noise (e.g. from atmospheric forcing), non-  
350 Gaussian forcing functions, strong non-linearities in ocean physics, and non-stationary behavior  
351 from external forcing and lack of equilibrium in the ocean.

352 Investigating the contribution of specific mechanisms to the predictability structures of  $\eta$   
353 is outside the scope of this study, but results presented suggest such analysis is merited. For  
354 example, work with linear models of  $\eta$ , includes that of Hughes and Williams (2010). Using  
355 altimetry alone, and without a prediction focus, they fit AR( $n$ ) models, choosing the orders using  
356 the Bayesian Information Criterion (BIC). With their higher temporal resolution, the AR( $n$ ) fits

357 are not easily compared to those from the use of the ARMA( $n, m$ ) on monthly and annual average  
358 values. They concluded that Rossby wave patterns are important within +/-30° of the equator,  
359 while advective processes become more influential at higher latitudes allowing features such as  
360 the Pacific-Antarctic ridge to influence  $\eta'$  predictability. The ARMA fits to the ECCOv4 data also  
361 have distinct features associated with the Pacific-Antarctic ridge and different behavior within  
362 30° of the equator, suggesting the physics that give rise to these features are coherent across these  
363 timescales and lend themselves to linear modeling approaches.

364

365 Predictability from annually averaged data as in figures 8a and 8b prove generally different.  
366 With the trend included, a region in the western South Pacific Ocean has striking performance.  
367 Removing the trend, an increase in prediction performance is seen in a band extending north-east  
368 of Hawaii, likely due to noise. Given the long time scales controlling oceanic physics, the records  
369 remain far too short to infer statistically stable results. In this context, the continuing difficulties  
370 generally experienced in distinguishing the lowest frequencies present between a general rednoise  
371 process and a true secular trend of multi-decadal applicability, remains a major issue. Whether  
372 unconstrained models, such as the CMIP5 ones used by Lyu *et al.* (2014), have true prediction  
373 skill remains unknown. Note too, that the univariate approach used here is readily extended to  
374 accommodate multivariate predictive models employing correlated spatial structures of many  
375 different types and which may work much more effectively in some areas.

376

377 In brief summary: the present study produces a benchmark of univariate linear skill in predicting  
378  $\eta$ . Figure 2b illustrates that up to 50% of the ocean is predictable using only the seasonal signal.  
379 The remaining ocean has an expected prediction error growth largely taking over two months to  
380 exceed 1 cm. Figure 6b shows that treating the linear trend as part of the continuum enhances

381 the predictive performance as expected. In areas in the Southern and Pacific Oceans, the stochas-  
382 tic continuum,  $\eta'$ , contains more than 95% of the total variance of  $\eta$ , with expected prediction  
383 performance of 1 cm exceeding a year in significant portions of these regions. Moving forwards,  
384 extending in time the global measurements of  $\eta$  and understanding the underlying physical pro-  
385 cesses remain the key to progress in regional sea surface height prediction.

386 **5. Acknowledgments**

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

390 **APPENDIX**

391 **Influence of chosen Information Criteria**

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

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

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

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

400

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

405

406     Figure 10 illustrates the orders chosen by the BIC for  $\eta'^\dagger$ . Note the difference to Figure 3 where  
407     higher orders are chosen using the AIC. Overall, the AIC produces estimates at higher orders than  
408     the BIC, but results for large areas are similar: e.g. the North Pacific and equatorial Atlantic. The  
409     information criteria, particularly the AIC, tend to pick out differing dynamical regions.

410

411     The predictive potential associated with the  $\eta'^\dagger$  using the BIC is illustrated in Figure 11. This  
412     shows a very similar pattern to Figure 6a, and we highlight the differences by showing the  
413     difference between the two (AIC-BIC) in Figure 12a. Here we see that the AIC largely gives  
414     higher predictive performance based on the ARMA( $n, m$ ) expected prediction error growth. The  
415     difference between the AIC and BIC prediction performance is highlighted in Figure 12b. We  
416     see slight differences, with the AIC having higher predictive performance to in bands  $+/-30^\circ$  of  
417     the Equator, particularly in the Pacific and Indian Ocean but also in the Atlantic. The BIC shows  
418     somewhat higher predictive performance in the higher latitudes, particularly in the Pacific.

419

420     For annually averaged  $\eta'$  the difference between using the AIC and BIC are smaller. Figure 12c  
421     highlights the difference between the AIC and BIC criterion, showing that overall the AIC has  
422     higher performance based on the ARMA( $n, m$ ) expected prediction error growth. With annually  
423     averaged  $\eta'^\dagger$ , the AIC and BIC also give similar prediction accuracies, with the differences  
424     highlighted in the difference plot shown in Figure 12d. However, small areas show higher  
425     performance based on ARMA( $n, m$ ) expected prediction error growth with the BIC.

426

427     The difference in predictability with the AIC and BIC informing the choice of the order suggests  
428     that the AIC has the highest utility. The BIC underestimates the order, and the AIC is found to be  
429     more suitable. This is illustrated in detail looking at the predictability with the apparent trend in  $\eta'$

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

446

## 447 References

- 448 Adcroft, A., Hill, C., Campin, J. M., Marshall, J., & Heimbach, P. (2004). Overview of the formulation and numerics of the MIT GCM (pp.  
449 139-150). Presented at the ECMWF Conference Proceedings, Shinfield Park, Reading, UK.
- 450 Aho, K.; Derryberry, D.; Peterson, T. (2014), "Model selection for ecologists: the worldviews of AIC and BIC", Ecology 95: 631-636,  
451 doi:10.1890/13-1452.1.
- 452 Akaike, H. (1973), "Information theory and an extension of the maximum likelihood principle", in Petrov, B.N. ; Csaki, F., 2nd International  
453 Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971, Budapest: Akademiai Kiad, p. 267-281.

- 454 Bingham, R. J., C. W. Hughes, V. Roussenov, and R. G. Williams (2007), Meridional coherence of the North Atlantic meridional overturning  
455 circulation, *Geophys. Res. Lett.*, 34, L23606, doi:10.1029/2007GL031731.
- 456 Box, G. and Jenkins, G. (1970). *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- 457 Burnham, K. P.; Anderson, D. R. (2004), "Multimodel inference: understanding AIC and BIC in Model Selection" (PDF), *Sociological Methods*  
458 & Research 33: 261-304, doi:10.1177/0049124104268644.
- 459 Burnham, K. P.; Anderson, D. R. (2002), *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach* (2nd ed.),  
460 Springer-Verlag, ISBN 0-387-95364-7.
- 461 Chelton, D. B., and M. G. Schlax, 1996: Global observations of oceanic Rossby waves. *Science*, vol. 272, pp. 234-238.
- 462 Chowdhury, MD. R., Chu, P.-S., Schroeder, T. and Colasacco, N. (2007), Seasonal sea-level forecasts by canonical correlation analysisan opera-  
463 tional scheme for the U.S.-affiliated Pacific Islands. *Int. J. Climatol.*, 27: 1389-1402. doi:10.1002/joc.1474
- 464 Cipollini, P., D. Cromwell, P. G. Challenor and S. Raffaglio, Rossby waves detected in global ocean colour data, *Geophysical Research Letters*,  
465 Vol. 28, No. 2 , pp. 323-326, 2001.
- 466 ECCO Consortium, 2017a, A Twenty-Year Dynamical Oceanic Climatology: 1994-2013. Part 1: Active Scalar Fields: Temperature, Salinity,  
467 Dynamic Topography, Mixed-Layer Depth, Bottom Pressure.
- 468 ECCO Consortium, 2017b, A Twenty-Year Dynamical Oceanic Climatology: 1994-2013. Part 2: Velocities, Property Transports, Meteorological  
469 Variables, Mixing Coefficients.
- 470 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  
471 modeling and global ocean state estimation, *Geo. Sci. Model Dev.*, 8, 2015
- 472 Forget, G. and Ponte, R.M., The partition of regional sea level variability, *Progress in Oceanography*, 137, 173-195, 2015.
- 473 Gille, S. T., 1994. Mean sea surface height of the Antarctic Circumpolar Current from Geosat data: Method and application, *J. Geophys. Res.*, 99,  
474 18,255-18,273.
- 475 Paul B. Goddard, Jianjun Yin, Stephen M. Griffies, Shaoqing Zhang. An extreme event of sea-level rise along the Northeast coast of North America  
476 in 2009-2010. *Nature Communications*, 2015; 6: 6346 DOI: 10.1038/ncomms7346
- 477 Hare, S. R., and N. J. Mantua, 2000: Empirical evidence for North Pacific regime shifts in 1977 and 1989. *Progress in Oceanography*, Vol. 47,  
478 Pergamon, 103-146.

- 479 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  
480 circulation on subannual to interannual timescales, *Ocean Sci.*, 9, 805-823, doi:10.5194/os-9-805-2013, 2013.
- 481 Chris W. Hughes, Simon D. P. Williams, The color of sea level: Importance of spatial variations in spectral shape for assessing the significance of  
482 trends, *Journal of Geophysical Research*, 2010, 115, C10
- 483 Church, J.A., et al., 2013: Sea Level Change. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the  
484 Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J.  
485 Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Chapter 13, pp. 1137-1216, Cambridge University Press.
- 486 Lin, X., J. Yang, D. Wu, and P. Zhai (2008), Explaining the global distribution of peak-spectrum variability of sea surface height, *Geophys. Res.*  
487 Lett., 35, L14602, doi:10.1029/2008GL034312.
- 488 Lorenz, E.N.: Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, 20 (2), 130-141, doi: 10.1175/1520-0469(1963)  
489 020<0130:DNF>2.0.CO;2, 1963.
- 490 Lorenz, E.N.: Low-order models and their uses, in G. N. Cathy Nicolis (ed.), Irreversible phenomena and dynamical systems analysis in geo-  
491 sciences, D. Reidel Publishing Company, pp. 557-567, 1987.
- 492 Lyu, K., Zhang, X., Church, J.A., Slanger, A.B.A., Hu, J., (2014). Time of emergence for regional sea-level change. *Nature Climate Change* 4,  
493 1006-1010. doi:10.1038/nclimate2397
- 494 J. V. Mecking, S. S. Drijfhout, L. C. Jackson, T. Graham (2016) Stable AMOC off state in an eddy-permitting coupled climate model. *Climate  
495 Dynamics*
- 496 Mengel, M., Levermann, A. (2014): Ice plug prevents irreversible discharge from East Antarctica. *Nature Climate Change* (online) [DOI:  
497 10.1038/NCLIMATE2226]
- 498 Melillo, J.M., T.C. Richmond and G.W. Yohe, Eds. (2014), Climate Change Impacts in the United States: The Third National Climate Assessment,  
499 U.S. Global Change Research Program, 841 pp., doi:10.7930/J0Z31WJ2
- 500 Milly, Paul CD, Kathryn A. Dunne, and Aldo V. Vecchia. "Global pattern of trends in streamflow and water availability in a changing climate."  
501 Nature 438.7066 (2005): 347-350.
- 502 Minobe, S., 1999: Resonance in bidecadal and pentadecadal climate oscillations over the North Pacific: Role in climate regime shifts. *Geophys.  
503 Res. Lett.*, 26, 855-858.

- 504 Ocaña, V., E. Zorita, and P. Heimbach (2016), Stochastic secular trends in sea level rise, *J. Geophys. Res. Oceans*, 121, 2183-2202,  
505 doi:10.1002/2015JC011301.
- 506 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*,  
507 12, 1542-1548.
- 508 Overland, J. E., J. M. Adams, and H. O. Mofjeld, 2000: Chaos in the North Pacific: Spatial modes and temporal irregularity. *Progress in Oceanog-*  
509 *raphy*, Vol. 47, Pergamon, 337-354.
- 510 Palmer, T.N., (2012) Towards the probabilistic Earth-system simulator: a vision for the future of climate and weather prediction. *Quarterly Journal*  
511 *of the Royal Meteorological Society* 138, 841-861.
- 512 "Sea Level as an Indicator of Climate and Global Change". Dr. Bruce Parker. *Marine Technology Society Journal*, Vol. 25, No. 4, 1991.
- 513 Donald B. Percival, James E. Overland, and Harold O. Mofjeld (2001) Interpretation of North Pacific Variability as a Short- and Long-Memory  
514 Process, *Journal of Climate*, 14:24, 4545-4559
- 515 Piecuch, C. G., and R. M. Ponte (2011), Mechanisms of interannual steric sea level variability, *Geophys. Res. Lett.*, 38, L15605,  
516 doi:10.1029/2011GL048440.
- 517 Pierce, D. W., 2001: Distinguishing coupled ocean-atmosphere interactions from background noise in the North Pacific. *Progress in Oceanography*,  
518 Pergamon, 49(1-4):331-352
- 519 Ponte, R.M. and C.G. Piecuch, 2014: Interannual Bottom Pressure Signals in the AustralianAntarctic and Bellingshausen Basins. *J. Phys.*  
520 *Oceanogr.*, 44, 14561465, <https://doi.org/10.1175/JPO-D-13-0223.1>
- 521 Priestley, M. B. (1981) *Spectral Analysis and Time Series*, London: Academic Press.
- 522 Rahmstorf, S., M. Perrette, and M. Vermeer (2012), Testing the robustness of semi-empirical sea level projections, *Climate Dynamics*, 39(3-4),  
523 861-875, doi: <http://dx.doi.org/10.1007/s00382-011-1226-7>.
- 524 Reynolds, R.W., D.B. Chelton, J. Roberts-Jones, M.J. Martin, D. Menemenlis, and C.J. Merchant, 2013: Objective Determination of Feature  
525 Resolution in Two Sea Surface Temperature Analyses. *J. Climate*, 26, 2514-2533, <https://doi.org/10.1175/JCLI-D-12-00787.1>
- 526 Roberts, C.D., D. Calvert, N. Dunstone, L. Hermanson, M.D. Palmer, and D. Smith, 2016: On the Drivers and Predictability of Seasonal-to-  
527 Interannual Variations in Regional Sea Level. *J. Climate*, 29, 7565-7585, <https://doi.org/10.1175/JCLI-D-15-0886.1>
- 528 Roemmich, D, Gould WJ, Gilson J. 2012. 135 years of global ocean warming between the Challenger expedition and the Argo Programme. *Nature*  
529 *Climate Change*. 2:425-428. 10.1038/nclimate1461

- 530 Thomson, D.J. 1994. Jackknifing multiple-window spectra.In: Pro-ceedings of the IEEE International Conference on Acoustics, Speech and Signal  
531 Processing, VI, 73-76.
- 532 Tulloch, R., J. Marshall, and K. S. Smith (2009), Interpretation of the propagation of surface altimetric observations in terms of planetary waves  
533 and geostrophic turbulence, *J. Geophys. Res.*, 114, C02005, doi:10.1029/2008JC005055.
- 534 Schott, A., Friedrich, P. and McCreary, J. (2001). The Monsoon Circulation of the Indian Ocean. *Progress in Oceanography - PROG OCEANOGR.*  
535 51. 1-123. 10.1016/S0079-6611(01)00083-0.
- 536 Shapiro, S.S. and Wilk, M.B. (1965). "An analysis of variance test for normality (complete samples)". *Biometrika*. 52 (3-4): 591-611.  
537 doi:10.1093/biomet/52.3-4.591. JSTOR 2333709. MR 205384. p. 593.
- 538 Stainforth, D., Aina, T., Christensen, C., Collins, M., Faull, N., Frame, D., Kettleborough, J., Knight, S., Martin, A., Murphy, J., Piani, C., Sexton,  
539 D., Smith, L., Spicer, R., Thorpe, A. and Allen, M.: 2005, Uncertainty in predictions of the climate response to rising levels of greenhouse  
540 gases, *Nature* 433(7024), 403-406. url: <http://oro.open.ac.uk/5025/>
- 541 D. Stammer, M. Balmaseda, P. Heimbach, A.Koehl, and A. Weaver, 2016: Ocean Data Assimilation in Support of Climate Applications: Status  
542 and Perspectives. *Ann. Rev. Mar. Sci.*, 8, 491-518.
- 543 Yang, Y. (2005), "Can the strengths of AIC and BIC be shared?", *Biometrika*, 92: 937-950, doi:10.1093/biomet/92.4.937.
- 544 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,  
545 5514?5520, doi:10.1002/2013GL057992.
- 546 Warren, Bruce A., Thomas Whitworth, and Joseph H. LaCasce. "Forced resonant undulation in the deep Mascarene Basin." *Deep Sea Research*  
547 Part II: Topical Studies in Oceanography 49.7 (2002): 1513-1526.
- 548 Wunsch, C., and P. Heimbach, 2007: Practical global oceanic state estimation. *Physica D*, 230, 197-208.
- 549 Wunsch, C. and P. Heimbach, 2013, Dynamically and kinematically consistent global ocean circulation and ice state estimates. In *Ocean Circulation*  
550 and Climate 2nd Edition, Siedler et al., Eds.
- 551 Wunsch, C. The interpretation of short climate records, with comments on the North Atlantic and Southern Oscillations. *Bull. Am. Met. Soc.* 80,  
552 245-255 (1999).
- 553 Wunsch, C. (2013), Covariances and linear predictability of the Atlantic Ocean, *Deep Sea Res., Part II*, 85, 228-243.
- 554 Wunsch, C. (2015), *Modern Observational Physical Oceanography: Understanding the Global Ocean*. Princeton University Press.

**555 LIST OF FIGURES**

**Fig. 1.** Example of the process of removing the seasonal signal at 67°S, 149°W. The mean (1992-2011) is removed, and the green line illustrates the fitted seasonal model ( $\bar{\eta}$ ), the blue line are the non-seasoned data ( $\eta'$ ), while the red line shows the original estimate from ECCOv4 ( $\eta$ ). . . . .

**Fig. 2.** From ECCOv4 of monthly averages from 1992-2011, Figure 2a shows the variance of  $\eta$  ( $m^2$ ), note the high variance associated with equatorial, western boundary currents and monsoonal regions. Figure 2b shows the percentage of variance contained in  $\bar{\eta}$  ( $\sigma_{\bar{\eta}}^2 / \sigma_{\eta}^2$ ). Figure 2c shows the percentage of variance contained in  $\eta'$  ( $\sigma_{\eta'}^2 / \sigma_{\eta}^2$ ). Contours correspond to 5% (black) and 95% (grey) of values. For reference, Figure 2d shows the amplitude of the variance ( $cm^2$ ) the  $\bar{\eta}$ , contributes towards the total variance of  $\eta$ . . . . .

**Fig. 3.** The chosen order of ARMA( $n, m$ ) using the AIC for the  $\eta'^{\dagger}$  1992-2011 (figures 3a and 3b showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively), and for non-seasoned  $\eta'$  data with the linear trend not removed (figures 3c and 3d showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively). . . . .

**Fig. 4.** Shapiro-Wilk test for normality for  $\eta'^{\dagger}$  (4a) and  $\eta'$  (4b). 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 are non-invertible. . . . .

**Fig. 5.** The prediction error as defined in equation 5 for  $\eta'^{\dagger}$  in three locations in the Southern Ocean (blue line, 50°S, 160°E), North Pacific (green line, 44°N, 150°E) and the Equatorial Pacific (red line, 2°S, 100°E). The rate of error growth varies regionally, showing the e-folding timescale (black bars). . . . .

**Fig. 6.** The ARMA( $n, m$ ) expected prediction performance of  $\eta'^{\dagger}$  (6a) and  $\eta'$  (6b) phrased in terms of the time it takes the expected prediction performance to be less than 1 cm (months). Areas saturating the colorscale indicate predictability longer than 17 months. . . . .

**Fig. 7.** The Shapiro-Wilk test for normality for the annually averaged  $\eta'^{\dagger}$  (7a) and  $\eta'$  (7b) 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 are non-invertible consistent with apparent non-stationarity, likely due to noise. . . . .

**Fig. 8.** The prediction performance based on the ARMA( $n, m$ ) expected prediction error for annually averaged and  $\eta'^{\dagger}$  (8a) and  $\eta'$  (8b). Performance is phrased in terms of the time it takes the prediction performance to be less than 1 cm (years). Areas saturating the colorscale indicate non-stationary ARMA coefficients. . . . .

**Fig. 9.** Chosen order of ARMA( $n, m$ ) using the AIC of non-seasoned  $\eta'^{\dagger}$  annually averaged with the linear trend removed (figures 9a and 9b showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively), and  $\eta'$  annually averaged with the linear trend not removed (Figures 9c and 9d showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively). . . . .



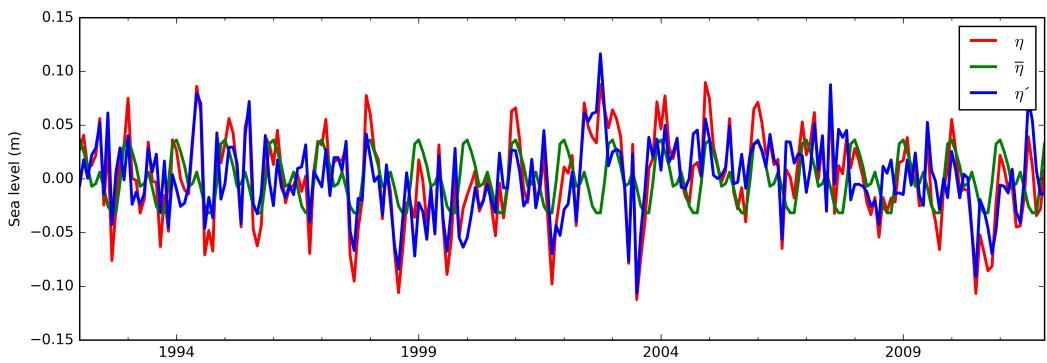


FIG. 1: Example of the process of removing the seasonal signal at  $67^{\circ}\text{S}$ ,  $149^{\circ}\text{W}$ . The mean (1992–2011) is removed, and the green line illustrates the fitted seasonal model ( $\bar{\eta}$ ), the blue line are the non-seasoned data ( $\eta'$ ), while the red line shows the original estimate from ECCOv4 ( $\eta$ ).

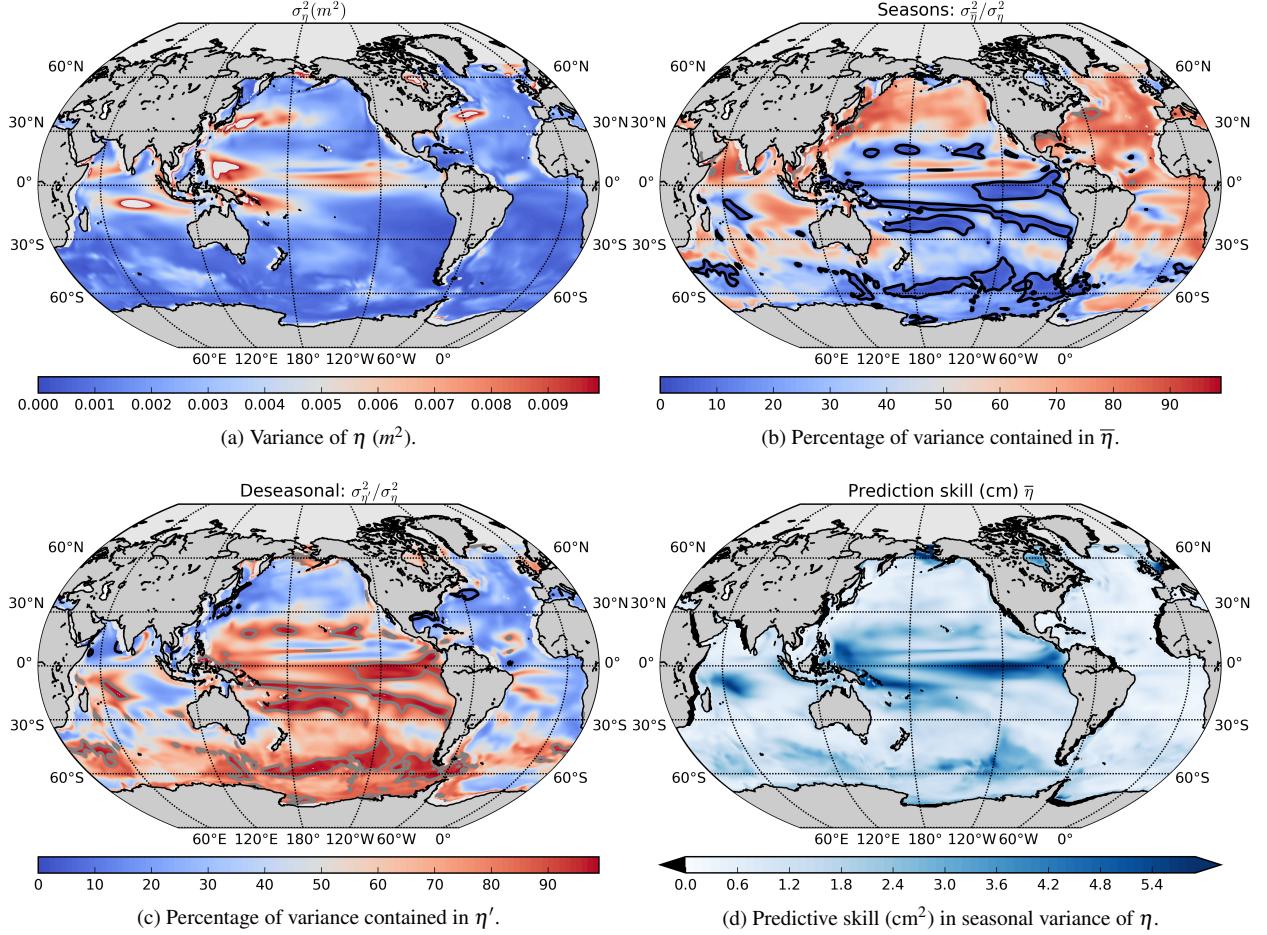


FIG. 2: From ECCOv4 of monthly averages from 1992-2011, Figure 2a shows the variance of  $\eta$  ( $m^2$ ), note the high variance associated with equatorial, western boundary currents and monsoonal regions. Figure 2b shows the percentage of variance contained in  $\bar{\eta}$  ( $\sigma_{\bar{\eta}}^2/\sigma_{\eta}^2$ ). Figure 2c shows the percentage of variance contained in  $\eta'$  ( $\sigma_{\eta'}^2/\sigma_{\eta}^2$ ). Contours correspond to 5% (black) and 95% (grey) of values. For reference, Figure 2d shows the amplitude of the variance ( $cm^2$ ) the  $\bar{\eta}$ , contributes towards the total variance of  $\eta$ .

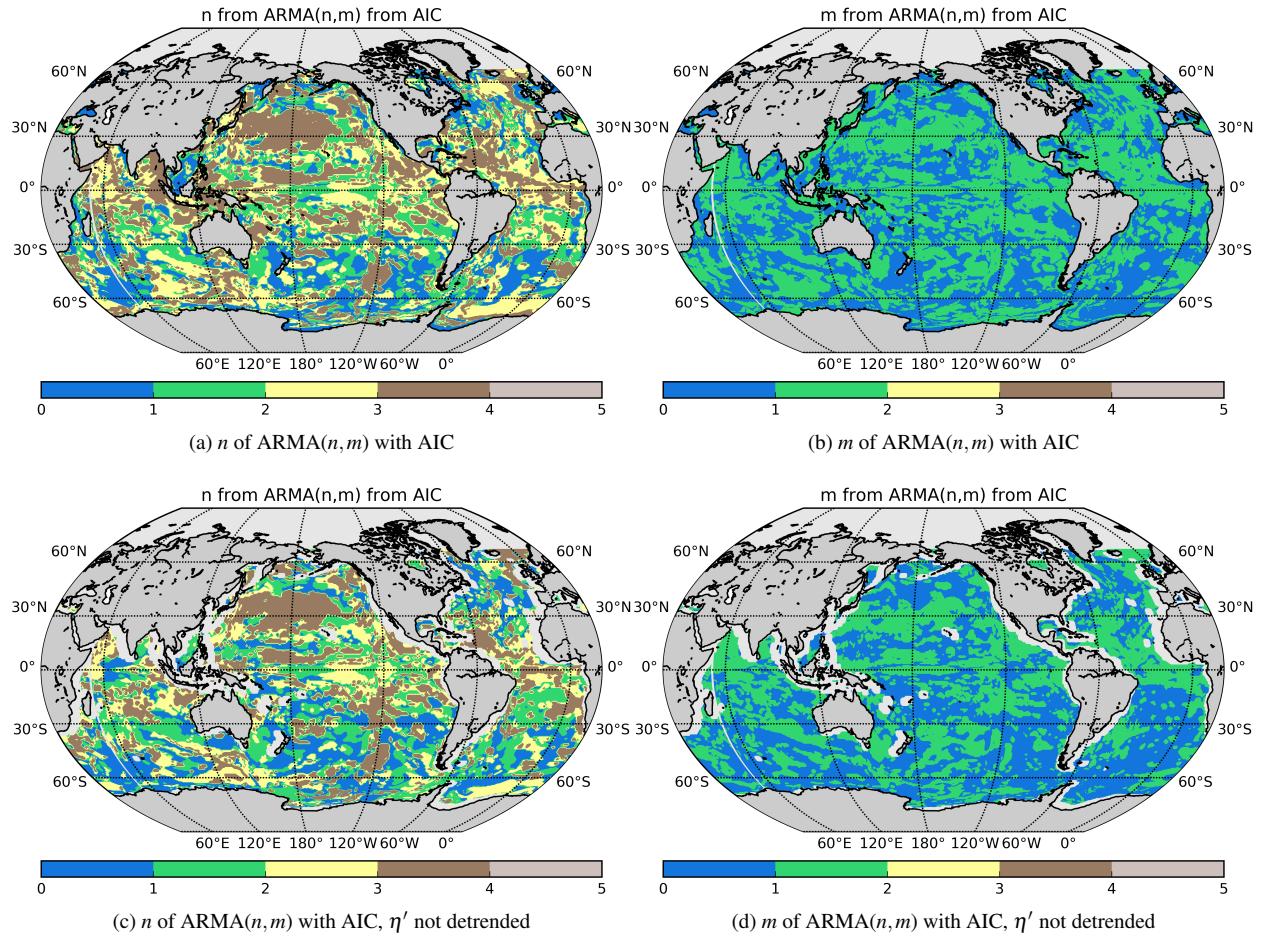


FIG. 3: The chosen order of ARMA( $n, m$ ) using the AIC for the  $\eta'^\dagger$  1992–2011 (figures 3a and 3b showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively), and for non-seasoned  $\eta'$  data with the linear trend not removed (figures 3c and 3d showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively).

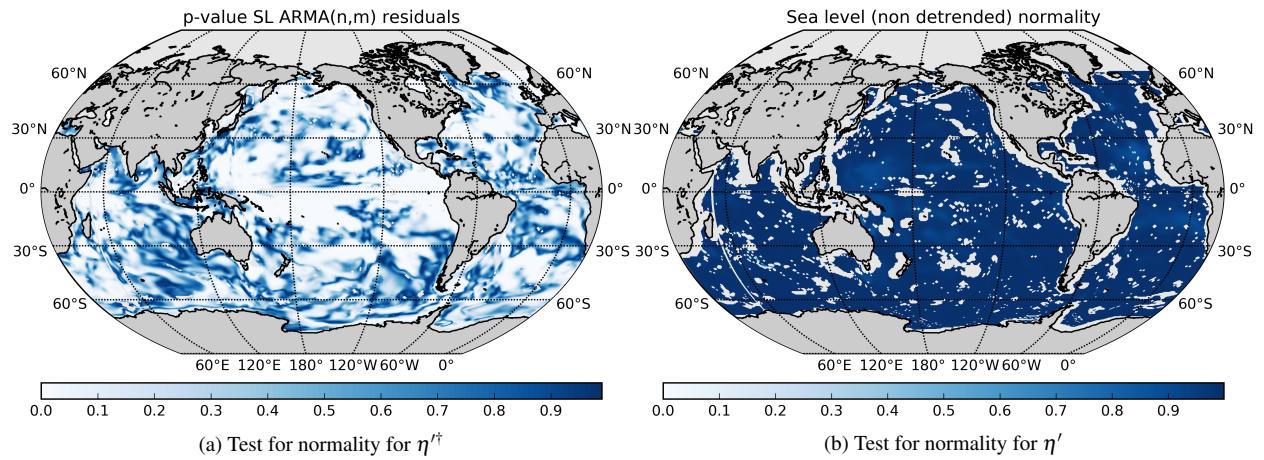


FIG. 4: Shapiro-Wilk test for normality for  $\eta'^\dagger$  (4a) and  $\eta'$  (4b). Darker blue indicating increasing confidence in accepting the null-hypothesis that the data is from a normally distributed population. White areas indicate areas where the AR coefficient roots were outside the unit circle, or where the MA coefficients were non-invertible.

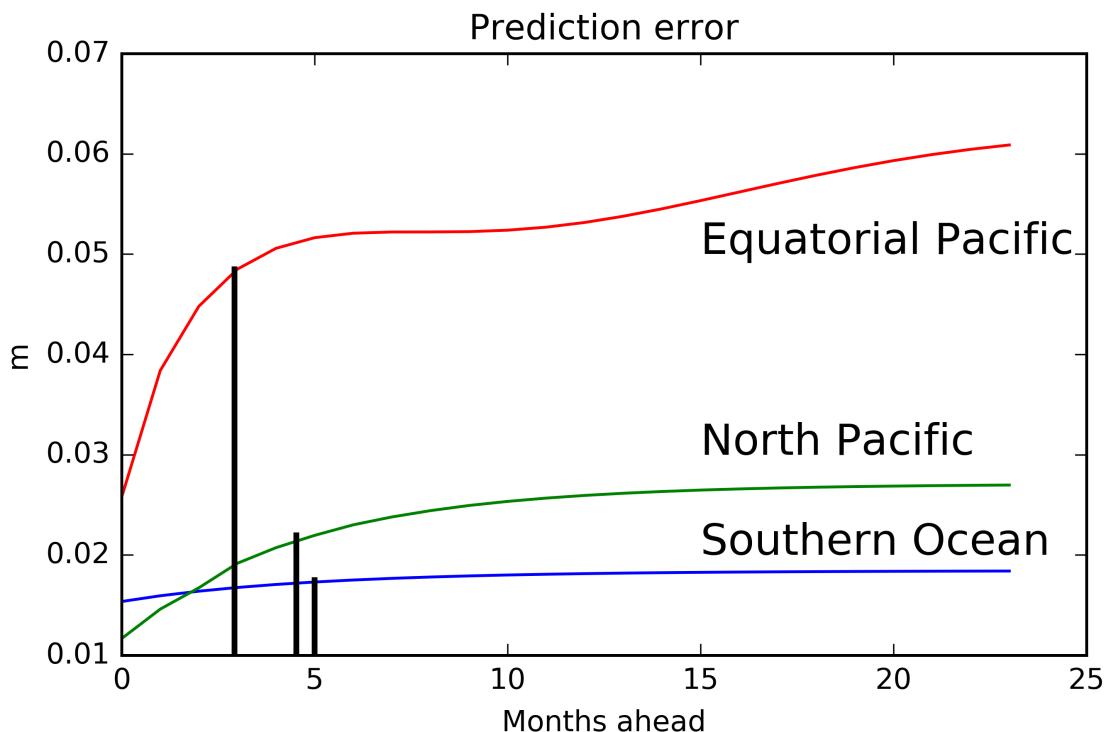


FIG. 5: The prediction error as defined in equation 5 for  $\eta'^\dagger$  in three locations in the Southern Ocean (blue line,  $50^\circ\text{S}$ ,  $160^\circ\text{E}$ ), North Pacific (green line,  $44^\circ\text{N}$ ,  $150^\circ\text{E}$ ) and the Equatorial Pacific (red line,  $2^\circ\text{S}$ ,  $100^\circ\text{E}$ ). The rate of error growth varies regionally, showing the e-folding timescale (black bars).

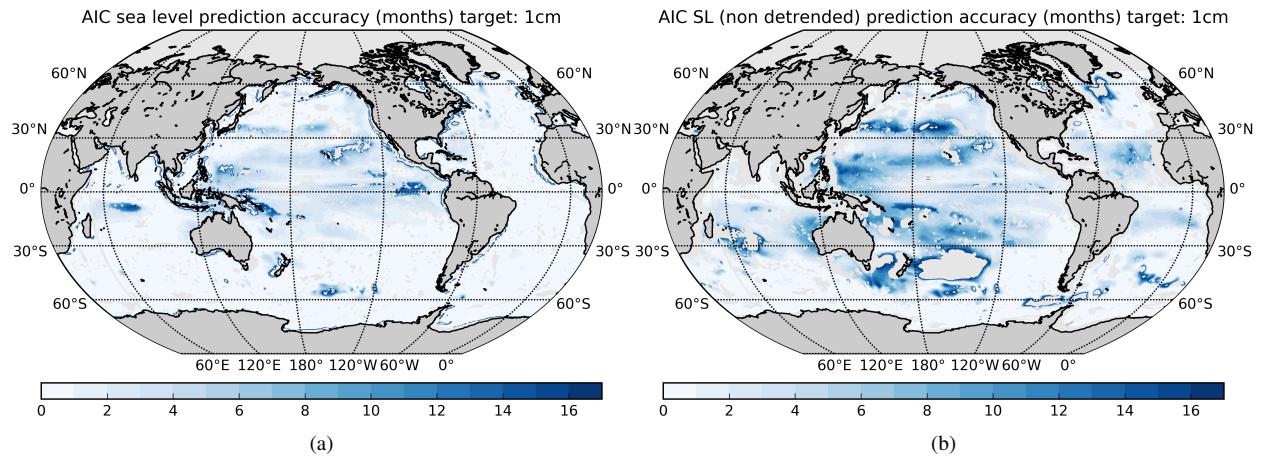


FIG. 6: The ARMA( $n, m$ ) expected prediction performance of  $\eta'^{\dagger}$  (6a) and  $\eta'$  (6b) phrased in terms of the time it takes the expected prediction performance to be less than 1 cm (months). Areas saturating the colorscale indicate predictability longer than 17 months.

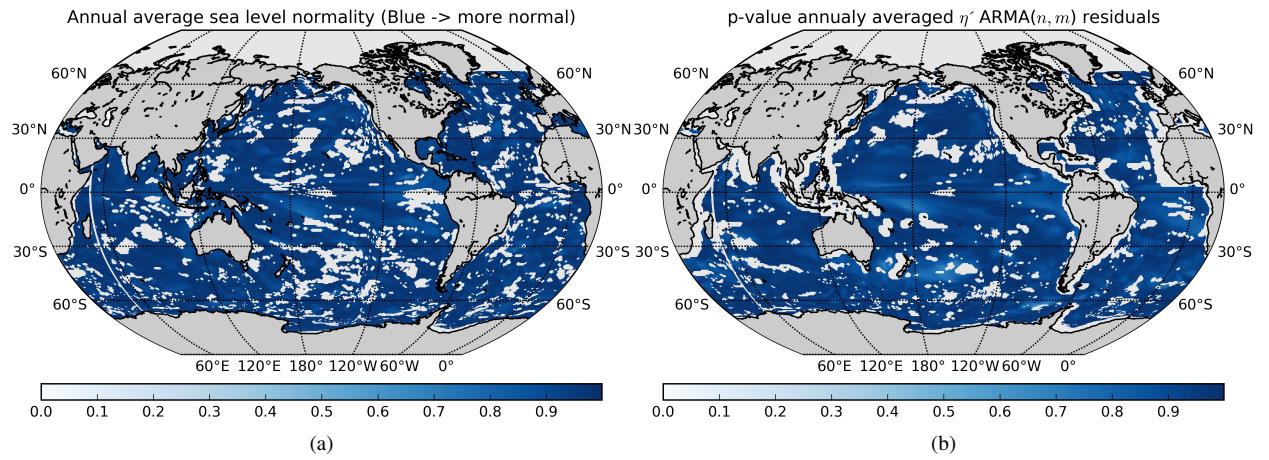


FIG. 7: The Shapiro-Wilk test for normality for the annually averaged  $\eta'^\dagger$  (7a) and  $\eta'$  (7b) data where the linear trend is not removed. Darker blue indicating increasing confidence in accepting the null-hypothesis that the data is from a normally distributed population. White areas indicate areas where the AR coefficient roots were outside the unit circle, or where the MA coefficients were non-invertible consistent with apparent non-stationarity, likely due to noise.

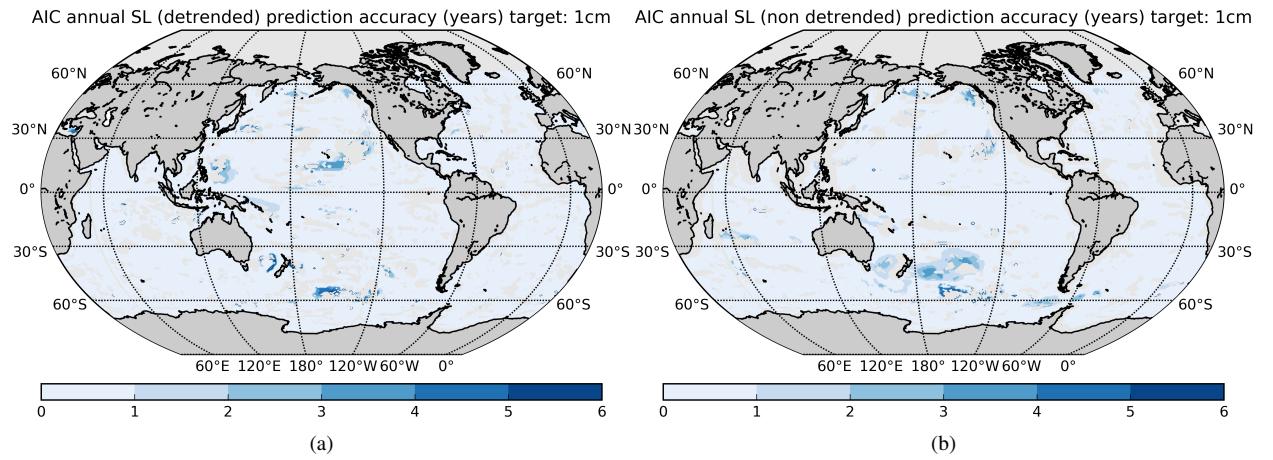


FIG. 8: The prediction performance based on the ARMA( $n, m$ ) expected prediction error for annually averaged and  $\eta'^{\dagger}$  (8a) and  $\eta'$  (8b). Performance is phrased in terms of the time it takes the prediction performance to be less than 1 cm (years). Areas saturating the colorscale indicate non-stationary ARMA coefficients.

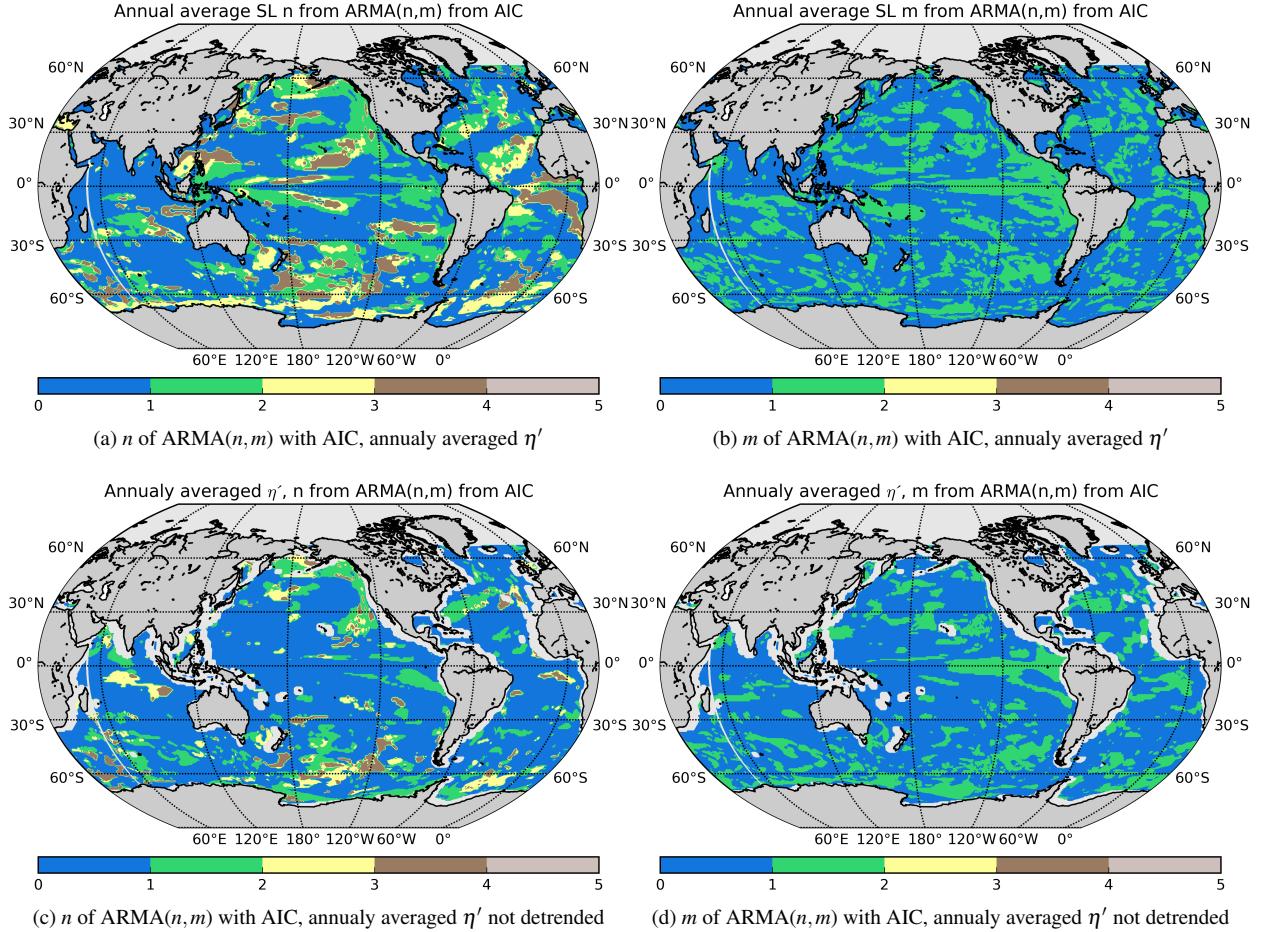


FIG. 9: Chosen order of ARMA( $n, m$ ) using the AIC of non-seasoned  $\eta'^\dagger$  annually averaged with the linear trend removed (figures 9a and 9b showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively), and  $\eta'$  annually averaged with the linear trend not removed (Figures 9c and 9d showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively).

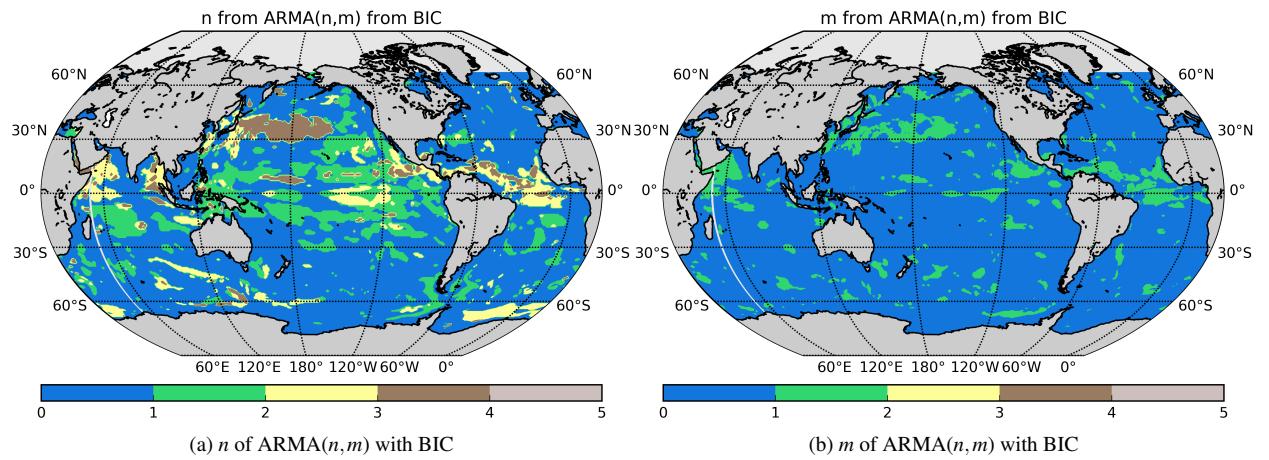


FIG. 10: The chosen order of ARMA( $n, m$ ) for annually averaged  $\eta'^d ag$  using the BIC (figures 10a and 10b showing the  $n$  and  $m$  of the ARMA( $n, m$ ), respectively). We use a four-point smoother for the  $\eta'^\dagger$  1992-2011. Note the difference to Figure 3 where higher orders are seen.

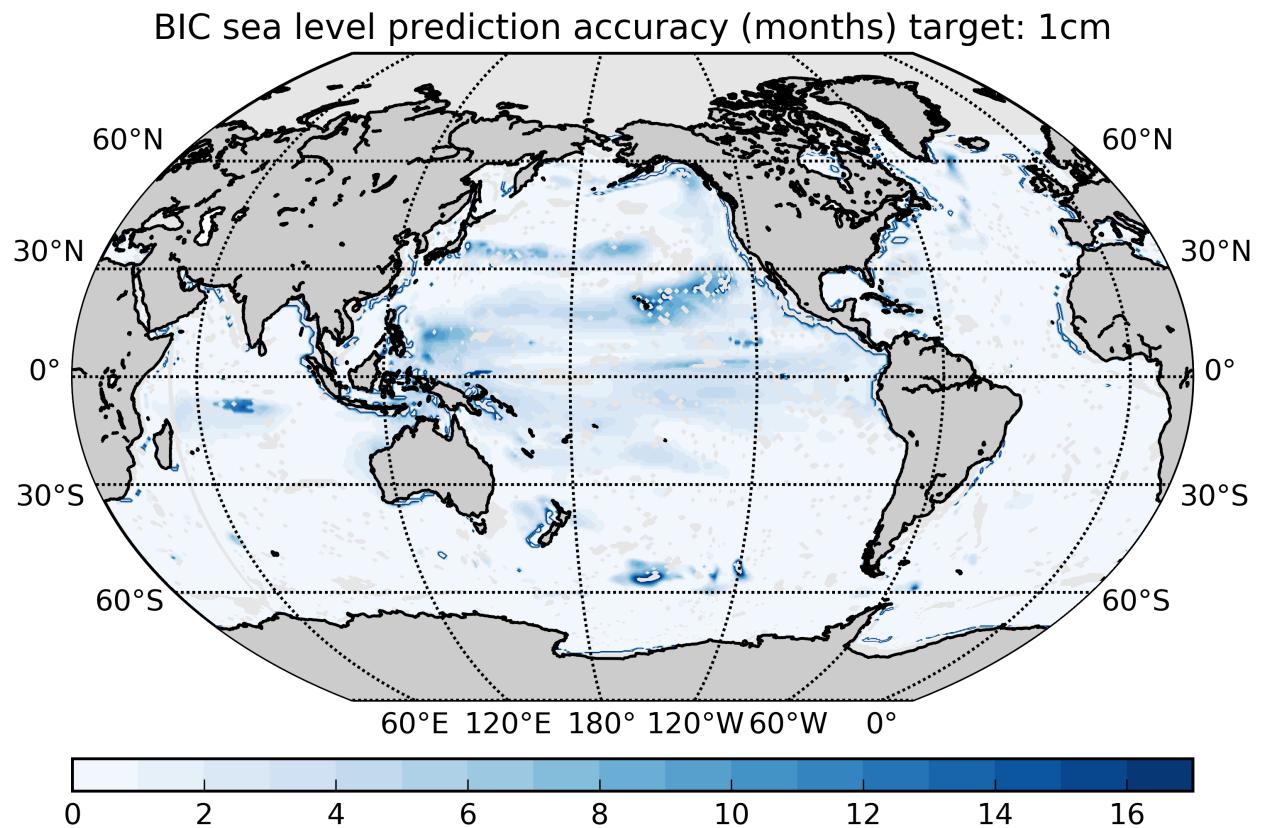


FIG. 11: The ARMA( $n, m$ ) expected prediction performance with BIC estimating the order of ARMA( $n, m$ ) models of  $\eta^{\dagger}$  phrased in terms of the time it takes the prediction performance to be less than 1 cm (months).

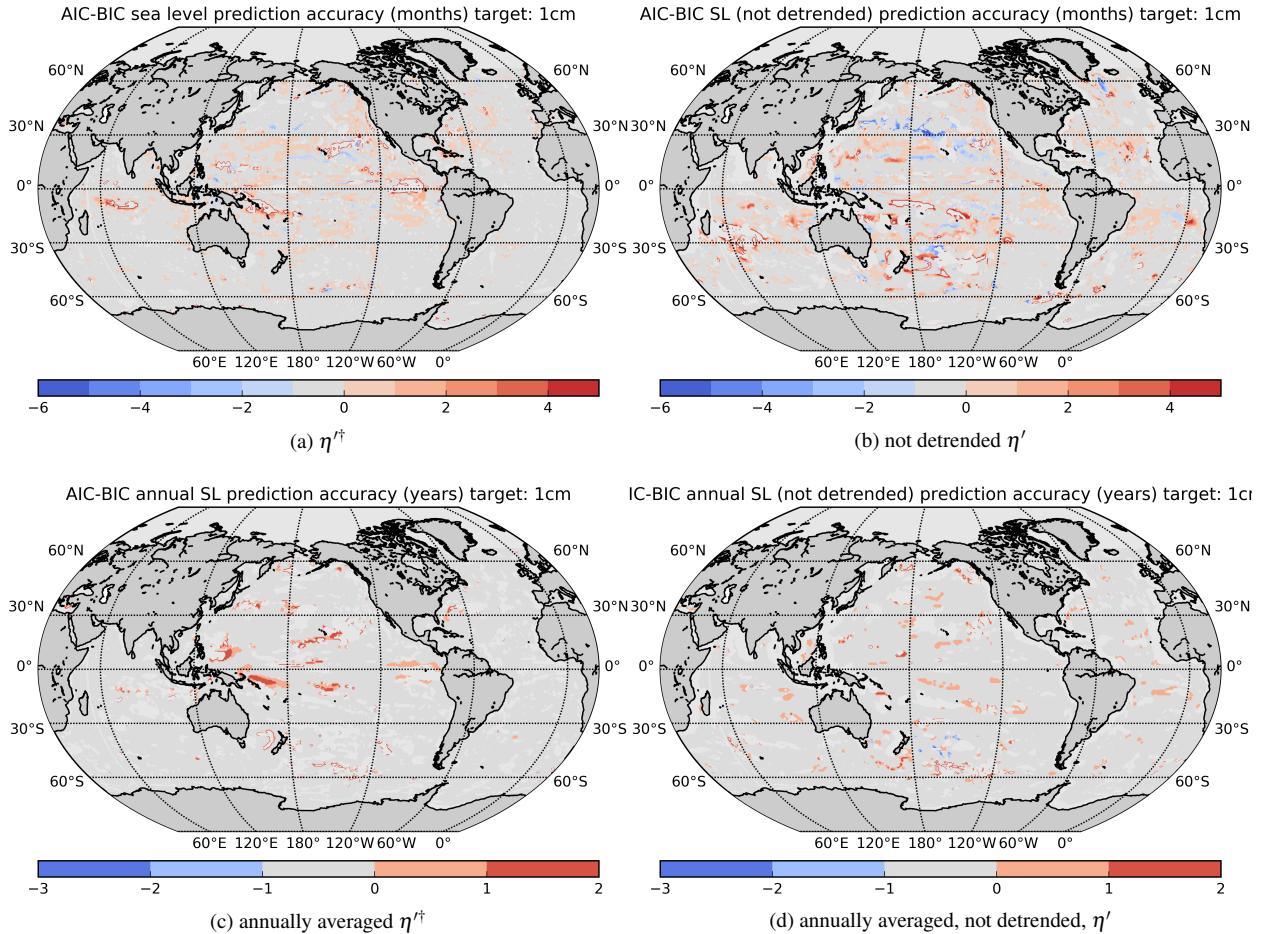


FIG. 12: The prediction performance difference with AIC-BIC using monthly averaged  $\eta'^\dagger$  (12a, in months), monthly averaged  $\eta'$  (12b, in months), annually averaged  $\eta'^\dagger$  (12c, in years) and annually averaged  $\eta'$  (12d, in years). Note that the AIC offers better predictability (red predominates) in most areas, but some prefer the BIC (blue).