

Covariances and Linear Predictability of the Atlantic Ocean

Carl Wunsch

Room 54-1426

Department of Earth, Atmospheric and Planetary Sciences,

Massachusetts Institute of Technology

Cambridge MA 02139 USA

email: cwunsch@mit.edu, Tel: 1-617-253-5937

August 23, 2011

Abstract

The problem of understanding linear predictability of elements of the ocean circulation is explored in the Atlantic Ocean for two disparate elements: (1) sea surface temperature (SST) under the storm track in a small region east of the Grand Banks and, (2) the meridional overturning circulation north of 30.5°S . To be worthwhile, any nonlinear method would need to exhibit greater skill, and so a rough baseline from which to judge more complex methods is the goal. A 16-year ocean state estimate is used, under the assumption that internal oceanic variability is dominating externally imposed changes. Linear predictability is the story of time and space correlations, and some predictive skill exists for a few months in SST, with some minor capability extending to a few years. Sixteen years is, however, *far too short* for an evaluation for interannual, much less decadal, variability, although orders of magnitude are likely stably estimated. The meridional structure of the meridional overturning circulation (MOC), defined as the time-varying vertical integral to the maximum meridional volume transport at each latitude, shows nearly complete decorrelation in the variability across about 35°N —the Gulf Stream system. If a time scale exists displaying coherence of the MOC between sub-polar and subtropical gyres, it lies beyond the existing observation duration, and that has consequences for observing system strategies and the more general problem of detectability of change.

1 Introduction

The ability to predict future climate is high on the agenda of many scientists (e.g., Meehl et al., 2009; Hurrell et al., 2010; Mehta et al., 2011). Claims that climate should be predictable

25 on some time-scale often rest upon the assumption that it would arise from the long memory of
26 the ocean—the atmosphere being assumed to lack such memory.

27 At the present time, more specifically, there is wide community interest in the possibility
28 of decadal prediction of some elements of the ocean circulation, including sea level changes
29 (e.g., Yin et al., 2009), surface temperatures (Newman, 2007), and volume transports (Zhang
30 and Wu, 2010; Msadek et al., 2010). Government funding agencies have issued calls for ac-
31 tual forecasts to be made (see e.g., the websites of the US National Science Foundation and
32 of the European Science Foundation). The comparatively short decadal time-scale raises the
33 possibility of observational tests of actual predictions, something that is implausible with 50
34 to 100 year forecasts—durations which exceed working scientific lifetimes, of model credibility,
35 and the interval since about 1992 of global-scale ocean observations. The extent, however, of
36 actual predictive skill for the ocean even on the decadal time-scale, much less the multi-decadal
37 one, remains obscure, with divergences of IPCC model extrapolations being a disquieting sign.
38 Some models are undoubtedly better than others, but which those are, and which fields are
39 well-calculated, remains unknown. Branstator and Teng (2010) review much of the existing
40 discussion.

41 Almost all studies of oceanic and its potential in climate predictability have been based upon
42 model calculations, and these have generally led to optimistic inferences (e.g., Msadek et al.,
43 2010). Some modelling studies have, however, led to more cautious conclusions. For example
44 Bingham et al. (2007) found little decadal meridional correlation between large-scale transport
45 characteristics—implying that any predictive skill in one region would have little impact on
46 larger scale, climatically important, components. In a study of the impact of noise disturbances
47 on the meridional overturning circulation (MOC), Zanna et al. (2011) found, for an idealized
48 configuration, that so-called non-normal error growth, particularly from small changes at depth
49 in sub-polar regions, would limit MOC predictive skill to considerably less than one decade.

50 In broader terms, predictability of the changes of any physical system involves several sub-
51 elements, including: the extent to which boundary conditions are predictable; the degree to
52 which variations arise from internal fluctuations with fixed or known boundary conditions; and
53 the degree to which that internal variability is fundamentally linear or non-linear. In particular,
54 any discussion of oceanic predictability confronts the awkward fact that the ocean tends to react,
55 rapidly and energetically, to shifts in the overlying atmosphere, particularly to changes in the
56 wind-field, most visibly in its upper reaches and often with little or no spatial correlation. (The
57 most rapid response is the barotropic one, which is almost instantaneous over the whole water
58 column.) A literature has emerged showing the coupling of the North Atlantic circulation to
59 the North Atlantic Oscillation (NAO, or Arctic Oscillation, AO) index; see e.g., Deser et al.

60 (2010). Some of the most important elements of the ocean circulation, as they affect climate,
61 such as the sea ice cover, or sea surface temperature (SST) are greatly modified by changing wind
62 systems, and they in turn, modify the atmosphere. This inference directs attention to the more
63 central question of whether the *atmosphere* is predictable on decadal time scales. No discussion
64 is provided of the probability that externally imposed finite amplitude shifts will occur, such as
65 the catastrophic collapse of the West Antarctic Ice Sheet (WAIS) and its numerous consequences.

66 The purpose of this paper is to explore some of the simpler aspects of the ocean prediction
67 problem employing, primarily, observations. The focus is on changes that are assumed, absent
68 strong evidence to the contrary, as arising from intrinsic ocean variability, rather than that
69 induced by global warming or other external drivers. Because there exist so many possible
70 predictable elements, we arbitrarily focus first on sea surface temperature (SST), and then on
71 the meridional overturning circulation (MOC) as exemplary of many of the issues. Attention
72 shifts to the most stable components embodied in the oceanic baroclinic structure. Simple theory
73 (Veronis and Stommel, 1956; Anderson et al., 1979) shows that, short of catastrophic external
74 disturbances, and outside of the equatorial band, basic characteristics such as the thermocline
75 depth and temperatures can be modified significantly only over many decades.

76 Notwithstanding several claims for the existence of major shifts in the ocean circulation,
77 there is *no* observational evidence in historical times of observed changes in basin-scale or larger
78 basic oceanic stratification or transport properties that lie beyond what are best labelled “per-
79 turbations” and for which linearization about a background state is a useful starting assumption.
80 One can compare e.g., the RRS Challenger (Tizard et al., 1873) hydrographic section, New York
81 to Puerto Rico, to recent sections nearby—without detecting any qualitative change. Rossby
82 et al. (2010) note that no detectable shift in mid-latitude Gulf Stream properties has occurred
83 over the last 80 years. It does remain possible that comparatively small changes in e.g., sea
84 surface temperature or sea ice cover, can generate major regional or global atmospheric climate
85 shifts—but if the oceanic component can be treated as essentially one of linear dynamics, a
86 substantially simplified oceanographic problem is the result.

87 The onset or suppression of such small spatial scale phenomena as rates, regions, and wa-
88 ter mass properties of convective regions are almost surely important to prediction skill over
89 long times as water mass production slowly accumulates. Convection and related processes
90 would generally have a nonlinear component—as they depend upon threshold-crossing physics.
91 Whether any existing nonlinear ocean model can reliably forecast such shifts is unknown. In
92 any case, Gebbie and Huybers (2011) show that surface sources of abyssal ocean waters are far
93 more widely distributed geographically than is conventionally believed.

94 If the perturbation depiction has any merit, it leads to the question of whether there is any

95 *linear* forecast skill. If the answer is “yes”, then any nonlinear approach e.g., through particle
96 filters, large ensembles, or simple runout of the of underlying GCM would have to exhibit a
97 significantly increased skill-level relative to the linear ones to justify the added expense. If the
98 answer is “no,” that there is no linear skill, one is led to understand how the nonlinear system
99 might be able, nonetheless, to produce a significant improvement. In any case, as for most
100 problems, it is worth exploring linear approximations before moving on to more complex forms.

101 Theoretical prediction skill is not meaningful unless it is coupled with a discussion of the
102 ability to detect it. Thus for example, a prediction that the meridional overturning circulation
103 will weaken by 1 Sv in 10 years might be correct, but if neither the present nor the future
104 values can be determined to that accuracy, at best one could say that the future value will
105 not be distinguishable from the present one. Observational detection accuracy is a function of
106 the scope and nature of the observation system, and of the structure of the variability noise in
107 the ocean. Although it is touched on only tangentially here and is rarely discussed elsewhere,
108 this issue of *detectability* is an essential ingredient in any useful discussion of forecast skill—and
109 deserves study in its own right. A closely related, also rarely discussed, question has already been
110 alluded to: what magnitude of change could be regarded as useful, for example, in producing a
111 measurable contribution to future climate shifts?

112 In proceeding, another difficult question concerns those elements one is trying to predict,
113 and why? Myriad choices are phenomenological (sea surface temperature, sea level, meridional
114 overturning,...), geographical (western North Atlantic, tropical eastern Pacific), seasonal (winter
115 time SST versus summer time), and time horizon (SST with a one month lead time can be of
116 intense interest to a weather forecaster, while the MOC state may be of interest only on 100+
117 year scales and then only to scientists). Here two fields of interest to different communities
118 (North Atlantic SST and the Atlantic MOC), are chosen, simplified as far as possible, and the
119 methodologies sketched that can be applied in seeking more definitive answers.

120 Linear predictability is the story of correlations of fields in space and time and thus their
121 estimates come to play the central role here. The observation-oriented approach, given the
122 extremely limited duration of large-scale oceanic observations relative to a multi-decadal re-
123 quirement, leads to the inference that one can hardly do more than state the problem. Resort
124 to models can and is being made, but the same data duration limitations preclude real model
125 tests.

126 2 An Ocean State Estimate

127 To proceed as best we can, the ocean state estimate ECCO-GODAE, v3.73, is used. This
128 estimate is discussed in detail by Wunsch and Heimbach (2007), Wunsch et al. (2009), and
129 in other papers listed on the website, <http://www.ecco-group.org>. For present purposes, a
130 sufficient description is that this state estimate is a near-global one over 16 years, from a least-
131 squares fit using Lagrange multipliers to the comparatively large oceanographic data sets that
132 became available beginning about 1992 in the World Ocean Circulation Experiment and later.
133 Adjustable parameters include initial conditions and all of the meteorological forcing functions.
134 The solution used is from this adjusted, and freely running, model. A partial discussion of the
135 time-mean of the estimate can be found in Wunsch (2011); the character of that mean relative
136 to dynamical equilibrium does have implications for predictability, and which will be touched
137 on at the end.

138 A terminology, “state estimate,” is used here to distinguish the result from estimates based
139 upon versions of meteorological forecast techniques (“data assimilation”)—which lead to prod-
140 ucts with physically impossible jumps and without global conservation principles. The results
141 here are primarily governed by observations, distinguishing them from the pure model runs:
142 Over the vast bulk of the oceans, the estimate is in a slowly time-evolving, volume and heat-
143 salt-conserving, thermal-wind balance, largely constrained by in situ hydrography, Argo float
144 profiles, and altimetric variability. It is thus a best-fit geostrophic, hydrostatic balance, in which
145 absolute velocities are determined from the conservation equations subject to Ekman pumping
146 and other surface forcing. Note that, among other data sets, monthly estimates of SST by
147 Reynolds and Smith (1995) were used.

148 Sixteen years is an extremely short period over which to determine multi-year or decadal
149 predictive skill. The restriction to that time period is dictated by the extreme paucity of oceanic
150 data prior to about 1992—when WOCE was underway. Ocean state estimates over intervals
151 before 1992 (e.g., Wang et al., 2010) are from nearly unconstrained ocean models. Furthermore,
152 the meteorological forcing fields used, even the most recent ones, have known major errors; see
153 e.g., Bengtsson et al. (2004) or Bromwich et al. (2007).

154 Because of the short-duration, a comparison will be made to the longer interval (28 years)
155 Reynolds and Smith (1995; hereafter RS) SST estimate used, separately, without the interven-
156 ing ECCO system. Such estimates are, however, not available for other fields of interest (the
157 meridional overturning, the corresponding oceanic heat transports, etc.), and for them the state
158 estimates must be used. The even-longer historical reconstructions of SST obtained prior to the
159 arrival of globally orbiting satellites are also avoided here, as the space-time sampling errors are

160 far worse.

161 **3 Sea Surface Temperature (SST)**

162 SST is always of central interest to meteorologists and provides a convenient starting point for
163 this investigation despite its being one of the most volatile and complex of all oceanic fields.
164 Vinogradova et al. (2010) discuss the global behavior of SST (particularly its rate of change)
165 in the ECCO solutions. Fig. 1 displays the time-mean SST over the 16-year duration of the
166 ECCO estimate

167 Woollings et al. (2010) have discussed elements of atmospheric storm track behavior re-
168 sulting from greatly increasing the SST resolution in the Gulf Stream region—where dominant
169 atmospheric cyclogenesis is thought to be most pronounced. The 1° version of the ECCO model
170 does not have sufficient resolution to reproduce the details of the Gulf Stream south of New
171 England, but it does do a reasonable job further north and east—in the sense of producing an
172 acceptable misfit to the data. Here the initial region of generic discussion is the small area
173 east of the Grand Banks depicted in Fig. 1, and which is close to being the eastern half of the
174 Woollings et al. (2010) region of interest. For the area (which will be referred to as the “Grand
175 Banks Box” or GBB, and denoted with a subscript G), the spatial average, $T_G(t)$, is formed
176 and is plotted in Fig. 2. The present focus on a small region contrasts with the notable effort
177 by Davis (1976) directed at the largest-scale features in the Pacific Ocean.

178 The time average of $T_G(t)$ is $\langle T_G(t) \rangle = 9.6 \pm 3.2^\circ\text{C}$. A simple, and perhaps even useful,
179 prediction of the temperature is its mean. In the present case, the annual cycle is so visually
180 apparent (not true of most oceanographic variables), that one is immediately led to a discussion
181 of its predictability. To the degree that it is purely periodic, one can extrapolate indefinitely into
182 the future. On the other hand, every seasonal cycle differs at least slightly from every other one,
183 and hence predictive skill will be imperfect. Fig. 3 displays the periodogram of $T_G(t)$, showing
184 that the annual cycle typically has about 90% of the variance over 16 years, with a smaller
185 contribution from the semiannual and higher harmonics. At this resolution, there is a sharp peak
186 at the annual period, of bandwidth less than the resolution limit of 1 cycle/16 years, meaning
187 that it is indistinguishable from a pure sinusoid. Note, however, that the background energy
188 surrounding and under this peak is not negligible and this energy prevents perfect prediction
189 of that component. (Methods exist, not necessary here, for predicting slowly changing annual
190 cycles; e.g., Hannan, 1970).

191 Using least-squares, the annual cycle and its first three harmonics were removed from the
192 record, leaving a residual, $T'_G(t)$, shown in Fig. 4, and producing an annual cycle amplitude

193 of $4.3 \pm 0.23^\circ\text{C}$ (the error is the formal one from the least-squares residuals). The variance of
194 the complete record is 10.2°C^2 , of which the deterministic annual cycle (and three overtones)
195 accounts for 9.4°C^2 or 93% (see Table 1). Variance dominance by the annual cycle is a challenge
196 to any model attempting to calculate either it, or the small deviations from it—should its
197 details change with climate. Of the residual 7%, most (about 5% of the total variance) lies in
198 periods longer than one year. Discussion of prediction now requires separating the problems at
199 interannual and intra-seasonal time scales.

200 3.1 A Formalism

201 With the removal of the annual cycle and its harmonics, as well as the time-mean, the residual
202 time series, $T'_G(t)$, can be assumed indistinguishable from a weakly stationary random process.¹
203 Many techniques exist for their prediction, and the literature is extremely large. Useful sum-
204 maries can be found in Robinson (1981), Hamilton (1994), Nelles (2001), Box et al. (2008),
205 von Storch and Zwiers (2001), and Priestley (1982, Ch. 10) among many others. General de-
206 velopments are associated with the names of Wold, Kolmogoroff, Wiener, Levinson etc., but
207 the most common formulation is through the development of autoregressive models of order
208 N (AR(N)), moving averages of order M (MA(M)), and combined models (ARMA(N, M)),
209 and their generalizations to non-stationary and nonlinear processes. Davis (1976, 1978, 1979)
210 provides excellent summaries of climate applications.

211 These linear methods, when new, were applied with a notable lack of success to ordinary
212 weather and stock market prediction. With understanding of the chaotic nature of weather,
213 the result is unsurprising. Rumors do persist that significant amounts of money can be made
214 using these methods in the stockmarket over minutes to hour time-scales, but on longer times
215 the stockmarket is not a stationary linear system. The present effort thus could be a quixotic
216 one—except that the degree to which, and which elements of the ocean circulation are chaotic
217 on decadal time scales, remains unknown. In any case, as argued above, there is little evidence
218 of large-scale deviations from slight perturbations in the observed circulation, and linearity is a
219 plausible starting point.

220 Here we will use primarily the AR and MA formulations (briefly summarized in Appendix
221 A) although the calculations are done in a slightly unorthodox manner to more directly empha-
222 size the underdetermined nature of the problem. Consider any zero-mean time series variable,
223 $\xi(t)$, which initially will be $T'_G(t)$. Suppose, to provide a specific example, that there exist L

¹Weak, or wide-sense, stationarity requires that the mean and second moments of the time series should be time-independent.

224 observations, including the present, and that it is an AR(2) process,

$$\xi(t) = a_1\xi(t-1) + a_2\xi(t-2) + \varepsilon(t), \quad (1) \quad \{\text{ar3}\}$$

where a_1, a_2 are unknown regression constants and $\varepsilon(t)$ is near-Gaussian white noise of zero mean and variance σ_ε^2 . Unless otherwise stipulated, t , denotes the present time, and the time-steps, Δt are implicit in all expressions. The coefficients in Eqs. (1) are in practice a set of simultaneous equations for the unknown $a_1, a_2, \varepsilon(r)$,

$$\begin{aligned} \xi(t) &= a_1\xi(t-1) + a_2\xi(t-2) + \varepsilon(t) \\ \xi(t-1) &= a_1\xi(t-2) + a_2\xi(t-3) + \varepsilon(t-1) \\ \xi(t-2) &= a_1\xi(t-3) + a_2\xi(t-4) + \varepsilon(t-2) \\ &\vdots \\ \xi(t-(L-3)) &= a_1\xi(t-(L-2)) + a_2\xi(t-(L-1)) + \varepsilon(t-(L-3)), \end{aligned} \quad (2) \quad \{\text{ar4}\}$$

for $L-2$ equations in L unknowns (a_1, a_2 , and $L-2$ of the $\varepsilon(r)$). Re-write Eq. (2) in standard matrix vector notation as,

$$\mathbf{E}\mathbf{x} = \mathbf{y}, \quad \mathbf{E} = \begin{Bmatrix} \xi(t-1) & \xi(t-2) & 1 & 0 & \cdot & 0 & 0 \\ \xi(t-2) & \xi(t-3) & 0 & 1 & \cdot & 0 & 0 \\ \cdot & \cdot & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi(t-(L-2)) & \xi(t-(L-1)) & 0 & 0 & \cdot & 0 & 1 \end{Bmatrix}, \quad (3) \quad \{\text{ls2}\}$$

$$\mathbf{x} = \begin{bmatrix} a_1 \\ a_2 \\ \varepsilon(t) \\ \varepsilon(t-1) \\ \cdot \\ \varepsilon(t-(L-3)) \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} \xi(t) \\ \xi(t-1) \\ \xi(t-2) \\ \cdot \\ \cdot \\ \xi(t-(L-3)) \end{bmatrix},$$

225 a formally underdetermined problem and which can be solved in numerous ways, including those
 226 commonly used in regression problems (e.g., Box et al., 2008; Priestley, 1982). The present
 227 formulation as a set of simultaneous equations differs from conventional least-squares (Priestley,
 228 1982, P. 346) only in treating the $\varepsilon(r)$ as explicitly part of the solution, rather than as residuals
 229 of the formally over-determined problem for a_1, a_2 alone. Here, for several reasons, we choose this
 230 depiction (Wunsch, 2006): the formal regression problem, when many more physical variables

231 are reasonably introduced (e.g., the SST time series at all latitudes, or the wind field), rapidly
 232 becomes very underdetermined even in the conventional formulation; least-squares makes simple
 233 the computation of uncertainties in the parameters $(a_1, a_2, \varepsilon(r))$; and one can easily “color” the
 234 noise $\varepsilon(t)$ either by modification of the identity matrix appearing in \mathbf{E} (which would make
 235 it an ARMA), or by introducing column weighting (solution covariance) matrices. Extension
 236 to arbitrary order AR processes is readily carried out. The normal equations governing the
 237 least-squares solutions of Eq. (3) involve the sample autocovariances of the ξ , known as the
 238 Yule-Walker equations.

239 For convenience in prediction, it is helpful to know that any stationary univariate AR can
 240 be converted into an MA, of form,

$$\xi(t) = \sum_{p=0}^{\infty} b_p \varepsilon(t-p) = \varepsilon(t) + b_1 \varepsilon(t-1) + b_2 \varepsilon(t-2) + \dots \quad (4) \quad \{\text{ma1}\}$$

241 For known a_i , the b_i can be obtained by algebraic long division,

$$1 + b_1 z + b_2 z^2 + \dots = \frac{1}{1 + a_1 z + a_2 z^2 + a_3 z^3 + \dots}, \quad (5) \quad \{\text{zpoly}\}$$

242 and vice-versa. The b_i can also be determined directly without first calculating the a_i . The MA
 243 form produces the τ -ahead *prediction error* (PE) as,

$$\left\langle \left(\tilde{\xi}(t+\tau) - \xi(t+\tau) \right)^2 \right\rangle = \sigma_\varepsilon^2 \sum_{p=0}^{\tau} b_p^2, \quad b_0 = 1, \quad (6) \quad \{\text{pe1}\}$$

244 the tilde denoting the prediction. This equation is obtained by substituting $\xi(t+\tau)$ into the
 245 left-hand-side of Eq. (6) and replacing the unknown and unpredictable $\varepsilon(t+1), \dots, \varepsilon(t+\tau)$ by
 246 their zero-means. If the b_i are sufficiently small, there will be rapid convergence to the asymptote
 247 of the variance of $\xi(t)$: $\langle \xi^2 \rangle = \sigma_\varepsilon^2 \sum_{p=0}^{\infty} b_p^2$. Like an N -order AR, any practical MA will have a
 248 finite order, M . Generally speaking if M is small, N will be large, and vice-versa, and with the
 249 trade-off becoming part of the discussion of representational efficiency. Note that stationarity,
 250 which we are assuming, requires that the polynomials in Eq. (5) should both be convergent
 251 when $|z| = 1$ (they are “minimum phase” in the signal processing terminology). Prediction
 252 error cannot exceed the variance of the time series—linear prediction cannot produce an error
 253 exceeding that from using the mean value.

254 Linear predictive skill for processes having a known power density spectrum can be deter-
 255 mined either by first computing the corresponding autocovariance and proceeding directly to the
 256 Yule-Walker equations, or more elegantly by using the Wiener-Kolmogoroff spectral factorization
 257 method (see Robinson, 1959, P. 105, or Priestley, 1982, Ch. 10). The spectral approach shows
 258 explicitly the connection between linear predictive power and the degree of frequency structure.

259 A time series with a flat (white) spectrum is unpredictable at any lead-time, τ , except for its
 260 mean value; structured spectra, including generic red noise, correspond to some additional linear
 261 predictive capability; and line spectra (pure periodicities) have infinite predictive time horizons.
 262 Many time series in nature are a mixture of these and other characteristics, and the fraction of
 263 the total variance that is predictable, and over what lead time, depends upon the details of the
 264 spectrum.

265 4 Months-Ahead Prediction

266 This autoregressive machinery is now used to estimate how predictable is $T'_G(t)$ (Fig. 4) about
 267 its mean, when sampled at monthly intervals? The red spectrum (an approximately -2.5 power
 268 law) of the residual (Fig. 3) shows that there is some predictability, dominated by the lowest
 269 frequencies. Because monthly and interannual physics are likely to be distinct, the question will
 270 be attempted in two stages: monthly mean samples and monthly forecasting and, annual mean
 271 samples and annual forecasting.

272 Because the solution to Eqs. (3) produces the same result as the conventional methods,
 273 standard statistical tests (e.g., Ljung, 1999; Priestley, 1982) can be used to infer that $T'_G(t)$
 274 can be represented as an autoregressive process with order between 3 and 6 (the tests differ).
 275 Because an AR(3) captures almost as much of the variance as do the higher order models, and
 276 is the simplest, we choose that as a reference case. The result, from solving the least-squares
 277 problem is,

$$T'_G(t+1) = 0.92(0.71)T(t) - 0.29(0.1)T(t-1) + 0.22(0.07)T(t-2) + \varepsilon(t+1), \quad \Delta t = 1\text{month},$$

278 where the parenthetical number is the standard error, with $\tilde{\sigma}_\varepsilon^2 = 0.2^\circ\text{C}^2$.

279 Directly estimating the MA form produces, alternatively,

$$T'_G(t) = 1.0\varepsilon(t) + 0.92\varepsilon(t-1) + 0.556\varepsilon(t-2) + 0.465\varepsilon(t-3) + 0.469\varepsilon(t-4) + \dots,$$

280 and which is slowly convergent. These MA forms were used to calculate the prediction error,
 281 which grows month-by-month (Fig. 5, Table 1) ultimately asymptoting after about 8 or 9 months
 282 to the full variance of $T'_G(t)$. (Recall that the total variance after removal of the annual cycle
 283 and its overtones is about 0.7°C^2 —and represents the maximum prediction error relative to the
 284 mean.) One might reasonably infer that there is useful (at the level of a few tenths of a degree
 285 error) linear predictive skill out to 4 or 5 months in the future, but not much beyond. Whether
 286 such skill is useful would depend upon the purpose of the prediction.

287 4.1 Comparison to the Satellite Record

288 Using the Reynolds and Smith (1995; RS) fields from this area, one can extend a similar SST
289 record out to 28 years. Details are not shown here, but a summary statement is that while the
290 monthly results differ in detail from those found for the ECCO-estimated record, there is no
291 qualitative difference, except that the apparent trend is more conspicuously reversing in recent
292 years (Fig. 6 and Table 1).

293 5 Interannual Behavior

294 Interannual behavior of the record is highly problematic: 16 samples (annual means) is far too
295 short to make much of any inference about correlation and prediction ability. The textbooks
296 already cited show how to calculate standard error statistics for the AR or MA coefficients, a_i, b_i ,
297 etc., and which depend directly on the autocovariances—assuming roughly Gaussian behavior.
298 To make the issue concrete, however, a small ensemble example for an AR(1)—the structure
299 with the fewest possible parameters other than white noise—is displayed in Appendix B and the
300 instability of the estimates from such small samples is obvious. We proceed here by restricting
301 the representation to an AR(1)—with the results interpreted cautiously as indicative only of
302 orders of magnitude.

303 5.1 Predicting Annual Averages of $T'_G(t)$

304 Fig. 6 shows the annual averages, $\bar{T}'_G(t)$, of the residuals of $T'_G(t)$ for both the state estimate and
305 the RS values. The state estimate shows a visible trend and a zero-order puzzle is the question of
306 whether that trend is a true secular one induced by global warming (defined here as extending
307 uniformly far beyond the record length), or a mere low frequency fluctuation manifested by
308 red noise (see Wunsch, 2010, for more discussion of the difficulties of trend determination, and
309 further references). Here it will arbitrarily be assumed that this signature is indeed a component
310 of red noise, as the longer RS record suggests, and thus will contribute to the predictive skill of
311 the interannual signal.

312 The one-year-ahead prediction error is approximately 0.03°C^2 rising to 0.2°C^2 after about
313 4 years (see Fig. 7 and Table 1). If a linear trend is first removed, neither the order nor
314 the prediction error (PE) are changed significantly. The RS results, not discussed, are very
315 similar. All that should be inferred is that linear predictive methods suggest some skill out to
316 about 5 years with errors of a few tenths of a degree. Whether any more sophisticated system
317 can do better remains, as of this writing, unknown.

318 **5.2 Predictability—A Caveat**

319 The reader is reminded that this study is based upon a “hindcast” skill, meaning that the same
320 data are used to determine the time series structure as are used to test its prediction skill.
321 Hindcast skill is inflated relative to true forecast skill by a significant amount. Davis (1976) has
322 a clear discussion of the issue. As he notes, an accurate estimate of the skill inflation is only
323 simple with large-sample statistics and, in particular, for interannual behavior, the estimated
324 SST used here is a very small sample. It is useful, in many cases, to withhold part of the data
325 set as a way of emulating an independent record for testing skill, perhaps by dividing it into two
326 pieces—an identification section and a test section. But the “red” nature of the spectra observed
327 shows that there will exist significant correlations between the used and withheld portions of
328 the time series, and again a rigorous calculation becomes difficult. We leave the discussion at
329 this point—as a warning that estimates here, particularly of the interannual forecast skill, are
330 optimistic ones.

331 **6 The Meridional Overturning Circulation (MOC)**

332 That the Atlantic MOC has become the center of so many studies, theoretical and observational,
333 is largely the result of the propagation of “conveyor belt” or “ribbon” pictures of the circulation,
334 whatever their physical reality might be. The MOC does provide a rough measure of the intensity
335 of the circulation in data and models. MOC connection to climate variability is, however, at best
336 indirect, and determining the volume or mass transport in the North Atlantic as a whole can be
337 done only by use of a model. A number of papers (e.g., Lorbacher et al., 2010) claim the existence
338 of useful covariances between MOC values and some observables such as sea surface height,
339 except these are also untested model results. Another immediate issue is the definition of what
340 is meant by the MOC, as the literature contains usages calculating it at very different latitudes,
341 integration depths, and averaging times. Here we take advantage of a global system to define
342 it—in the Atlantic Ocean—as a function of all latitudes from the Cape of Good Hope (about
343 30°S) northward to the northern limits of the present model (79.5°N). It is, more specifically,
344 calculated as the zonal integral at monthly intervals, continent to continent, of the meridional
345 velocity, the density being treated as constant, consistent here with the Boussinesq version of
346 the model,

$$V(y, z, t) = \int_0^{x_L(y)} v(x, y, z, t) dx \quad (7) \quad \{\text{meridtrans1}\}$$

347 (in practice, spherical coordinates are used). At any latitude, at any time, the MOC is then
348 arbitrarily defined as the maximum of the integral from the surface to a time and space varying

349 depth $z_{\max}(y)$,

$$V_{moc}(y, t) = \max_{z_{\max}(y, t)} \int_{z_{\max}(y, t)}^0 V(y, z, t) dz \quad (8)$$

350 Fig. 8 displays the time average value, $\langle V_{moc}(y, t) \rangle$ as well as the depth, z_{\max} , where, on
351 average, the maximum is reached (Fig. 9). A geographical maximum of about 16Sv is reached
352 at northern mid-latitudes and drops rapidly with latitude beyond about 50°N. At the present
353 time, it is not possible to provide a useful uncertainty estimate for these values, but the general
354 structure—mass-conserving thermal wind-balance—appears very robust to both variations in
355 the data base and in model parameters. The meridional flows, V , were discussed in some detail
356 by Wunsch and Heimbach (2006, 2009).

357 How much does $V(y, z, t)$ vary with time? Jayne and Marotzke (2001) infer, consistent
358 with what is found here, that the seasonal volume variability arises primarily in the surface
359 Ekman layer. Fig. 10 shows the meridional transport January anomaly values every two years,
360 indicating variations of up to about 4Sv, but only very locally—mainly in the vicinity of the
361 equator, and at about 40°N. The variations in the anomaly of $V_{moc}(y, t)$ are shown in Fig. 11 at
362 three latitudes, where the integration depth is kept fixed at $z_{moc}(y)$, that is not time-varying.
363 These integrals have a range, except in the far north, of about ± 5 Sv and are noisy on monthly
364 time scales. Temporal variances of V at all latitudes are depicted in Fig. 12. The power densities
365 for three latitudes are shown in Fig. 13. At most latitudes, there is a significant annual cycle and
366 its harmonics, especially in the low-latitude Ekman layer. Otherwise, the spectral densities are
367 nearly white beyond the annual period—boding ill for decadal linear predictability. The smallest
368 low frequency energy is found at 50.5°N, a result consistent with the linear dynamical behavior
369 there requiring much longer adjustment times. High latitude power densities are dominated
370 by the annual cycle and not by the interannual variability (out to 16 years). In general, these
371 spectra are “flat” by geophysical standards, being not very far from white noise.

372 Variances of the MOC, computed for the monthly means over all 111 latitudes are $27\text{Sv}^2 =$
373 $(5.1\text{Sv})^2$ and the annual means have variance $1.5\text{Sv}^2 = (1.2\text{Sv})^2$ providing a rough idea of
374 the temporal variability and the observational challenge. At 50.5°N alone, the corresponding
375 variances are $10.5\text{Sv}^2 = (3.2\text{Sv})^2$, and $1.8\text{Sv}^2 = (1.3\text{Sv})^2$.

376 A small visible trend appears early on in the values at some latitudes, a trend which disap-
377 pears as one moves away from the starting time. No data precede the start time of 1992; hence
378 the early years are much more weakly constrained than the later ones—which are controlled in
379 considerable part by the data preceding the particular time of estimation.

380 Fig. 14 shows the correlation coefficient matrix, R_{ij} , between the annual mean variations
381 in the MOC at all latitudes, i, j . Making the mildly optimistic assumption that each of the

382 annual mean values is an independent variable at any latitude, at 95% confidence, one must
383 have $|R_{ij}| > 0.5$, approximately, to distinguish the value from zero. A change takes place across
384 about 35.5°N where all linear correlation is lost between values on either side of that latitude
385 (the approximate Gulf Stream position). The North Atlantic subtropical gyre shows some
386 marginally significant correlation with the South Atlantic, but no correlation with the North
387 Atlantic subpolar gyre (consistent e.g., with the pure model results of Bingham et al., 2007)
388 except for a slight hint of a finite relationship between 75°N and the South Atlantic. Within
389 the subtropical gyre, correlation decays to insignificant levels beyond separations of about 20° of
390 latitude. (A more elaborate analysis by E. Haam, personal communication, 2011, using a Monte
391 Carlo simulation (Haam and Huybers, 2010), suggests that the small band of higher correlation
392 between about 75°N and the South Atlantic, visible as horizontal and vertical stripes in Fig. 14,
393 is statistically significant. An oceanic physical mechanism for “skipping over” the intermediate
394 latitudes is not obvious, and one probably must look to the atmosphere for an explanation.)

395 A problem with correlation analyses is that they lump together all time scales, often having
396 very diverse physics. One might hypothesize that the low correlations found here are the result
397 of noisy high frequencies. To address this issue in part, Figs. 15 and 16 show the coherence as
398 a function of frequency between the 50.5°N MOC and its values at 25.5°N and 20.5°S . They
399 show, to the contrary, that the only marginal coherence is at periods shorter than one year (at
400 the annual period the conventional statistics do not apply as sinusoids are always coherent).
401 Evidently (on this decadal time scale), annual mean MOC determinations south of about 35°N
402 carry no (linear) information about its behavior poleward of that latitude at any frequency now
403 testable.

404 One could search the system for correlations. For example, it is conceivable that there is
405 correlation between the meridional transports lying between some pair of isopycnals, at two
406 different latitudes, even though the total transport shows nothing significant. If one searches a
407 large number of possible combinations, some apparently significant relationship will necessarily
408 be found. If there are 100 possible combinations, then using a 95% level-of-no-significance with
409 proper probability densities, about 5% should show apparent, but spurious, correlation. This
410 direction is not pursued.

411 **6.1 Predicting the MOC**

412 Monthly predictions of the MOC have no obvious utility and they are not discussed here; only the
413 annual means are now considered. Wunsch and Heimbach (2009) discuss the annual cycle of the
414 MOC—and which is primarily a near-equatorial phenomenon, extending to considerable depth.
415 Hypothetically, one could imagine using each of the 111 time series at 1° latitude spacing, with

416 time lags of one year and longer, as regression variables to predict e.g., the value at some specific
417 latitude(s). With 16 sample points at any fixed latitude, one would be seeking the equivalent of
418 the expansion of a 16-dimensional vector in 111 non-orthogonal vectors—as in Appendix C—a
419 markedly underdetermined problem. Although we will return to this problem, consider instead
420 the more well-determined one of predicting from the present and past values at one particular
421 latitude. As for SST, the main problem is having only 16 samples,

422 The MOC at 50.5°N is arbitrarily chosen as the initial target prediction—on the basis of
423 a large literature claiming that modifications in the high latitude transports are a key climate
424 control parameter. This latitude is close to the one with the largest defined MOC and is just
425 south of the region where the mean MOC declines very rapidly. Thus consider the problem of
426 predicting the MOC at 50.5°N one year into the future, using the calculated history at that
427 latitude. The spectral estimate in Fig. 13 is not very different from white noise at long periods,
428 and one anticipates only some modest degree of prediction skill. Fig. 17 shows the error growth
429 using an AR(1) deduced from the measurements at 50.5°N alone (and see Table 1).

430 Had there appeared significant correlations or coherences between 50.5°N and other lati-
431 tudes, it would be reasonable to seek predictive power from observed variations in the MOC
432 at all latitudes. The absence of such correlations shows that linear predictability will be slight.
433 Experiments using singular vectors (not shown; Appendix C describes them), as expected did
434 not produce any useful outcome.

435 It is, of course, possible that the existing 16-year interval is untypical of the longer-term
436 behavior of the Atlantic Ocean and/or that linear predictive skill would emerge with much longer,
437 multi-decadal or centenary, records, but these are purely speculative claims. The utility for
438 prediction from existing duration, geographically widely separated, field observations is doubtful.

439 **6.2 Correlation with SST**

440 Study of the MOC has often been justified on the basis that its variability is linked to climate
441 change, sometimes in truly dramatic fashion (“hosing” and “shut-down”). Thus the question
442 arises as to whether there is any relationship between the MOC variations estimated here, and
443 the SST of the region previously discussed. One simple measure is the correlation coefficient
444 between the MOC and GBB SST variations, depicted in Fig. 18, which repeats Fig. 14, but such
445 that the last row and column now represent the annual mean SST time series. The calculation is
446 shown for the case of the raw SST and where, also, its visible, linear, trend was removed by least-
447 squares. One might infer that there is a marginally significant negative correlation between the
448 low latitude MOC ($0 \pm 10^\circ$ latitude) and the GBB SST. The result is, however, dependent upon
449 the presence of the trend in SST, and which destroys the assumption of annually independent

450 changes. Any inference of correlation is extremely fragile and not supportive of a relationship
451 between MOC and SST on the time scales accessible here. Determining whether there is such a
452 relationship on much longer time scales will have to wait on extended observations.

453 **7 Discussion**

454 Linear predictability in space-time systems is the story of the covariance structure. It is thus
455 useful to compare the results for the MOC here with the entirely different approach and inferences
456 of Msadek et al. (2010) as an example of a pure model approach. They concluded that the MOC
457 is predictable with some skill out to 20 years, using an unconstrained, coupled climate model
458 run for 1600 years. Apart from the very much longer analysis time, their mean MOC is 25 Sv
459 rather than the approximate maximum of 16 Sv found here. Their MOC spectrum (their Fig.
460 1) is steeply red from about two year periods to about 20 years, culminating in a narrow-band
461 spectral peak near 20 years. That their inferred predictability is larger than found here, at
462 about 20 years, would be a consequence of their narrow spectral peak at that period—if it is
463 real. This prediction skill is likely primarily a linear one, because low frequency narrow-band
464 processes have an intrinsic long memory—extended correlation times; as the peak-width becomes
465 narrower, one converges to a deterministic component with an infinite prediction horizon. In
466 contrast, the spectra computed here tend to indicate a white noise behavior beyond about
467 15-year periods with no indication of a narrow band spectral process, although no definitive
468 statement can be made from the available observations.

469 This disagreement between the two sets of results focuses one on the central conundrum of
470 climate change studies: (1) It is difficult to compare a 16-year data-constrained estimate to a
471 1600-year unconstrained one. (In their study of 136 years of North Atlantic SST data, Tourre
472 et al. (1999) did not report any obvious 20 year spectral excess, although all the caveats about
473 data quality before the polar-orbiting satellite era will apply, and even that recent system is
474 imperfect.) Conceivably, the present 16-year interval of the ECCO estimates is unrepresentative
475 e.g., of the historical strength of the MOC, and one might postulate that it was more typically
476 closer to the 25Sv of the Msadek et al. (2010) model than to the ECCO values of the WOCE era.
477 Such an enhanced value, however, would imply a much increased geostrophic transport, which
478 dominates the upper limb of the MOC —mostly in the Gulf Stream system. Within historical
479 times, such a large strength is probably ruled out by existing coastal sea level and wind-strength
480 records, but no quantitative estimate has been made. (2) Conceivably the more nearly white
481 spectrum that we infer at periods of a few years is also untypical of a hypothetical much longer
482 record. How does one know?

483 A general comment, applicable also to the present results, is that most models are much
484 less noisy than is the real world, either entirely lacking in the eddy field and internal waves,
485 or commonly underestimating them. In the present case (e.g., Wunsch, 2008; Kanzow et al.,
486 2009) and in calculations such as Msadek et al. (2010), one should infer that all estimates of
487 predictability (or its relative, detectability) skill are probably upper bounds.

488 Poor prediction results in the fields discussed here does *not* mean that the corresponding
489 variable is not predictable: sometimes the best prediction is just the sample mean, with a stan-
490 dard error given from the variance of the variable. That is, given observing system limitations,
491 and the great oceanic noisiness, the best prediction may well be that the field will be indistin-
492 guishable from present values—and that estimate may still be a useful one. A nonlinear method,
493 one that was independent of any linear space-time correlation, might well do better, although
494 the nonlinearity would have to be one operating on statistical moments higher than the second.
495 Note that methods exist for transforming some nonlinear time series into linear forms (e.g.,
496 Hamilton, 1994, etc.).

497 One can modify and extend the methods here in a large number of ways. The singular value
498 decomposition (see Appendix C) is identical in its \mathbf{u} vectors to the conventionally defined EOFs,
499 and emerges naturally as part of the least-squares/regression problem. These individual orthog-
500 onal structures of the variability have been used by Davis (1978) and many others. Generally
501 speaking, any particular EOF (singular vector) will have a fraction, depending upon the degree
502 of spatial correlation, of the total variance, and if it displays significant predictability (e.g.,
503 Branstator and Teng, 2010), it will only be for that fraction of the expected variance—perhaps
504 large enough to be useful to someone if its skill can be tested.

505 The dual (adjoint) model calculations of Heimbach et al. (2011) represent a running lin-
506 earization of the governing equations about the time-varying state. Regarded as Green function
507 solutions, they can be used either directly in predictions, or as a guide in choosing the relevant
508 regressor fields, locations, and time-scales. They do show the strong sensitivity of North At-
509 lantic shifts to disturbances in distant ocean basins at earlier times. On time scales of decades
510 and longer, variability in the Atlantic is a summation of disturbances emanating from the en-
511 tire global ocean. No single region dominates the later changes in the North Atlantic, and for
512 understanding and prediction, a global, long-duration observing system is required.

513 As noted in the Introduction, the present results apply only to the temporally statistically
514 stationary components. A major shift in the controlling boundary conditions—such as a massive
515 ice melt event, or an increase in greenhouse gases—would render the process non-stationary—
516 changing its mean, and likely its higher statistical moments as well. The issue for those inter-
517 ested in decadal and longer predictability is whether those external controls are predictable and

518 whether they dominate the variance contributed by what here is assumed to be intrinsic changes
519 in the ocean. Such external predictability, if it exists, is primarily independent of purely oceanic
520 processes and their long memory components. A long memory has the consequence, however,
521 of producing changes today or in the future as the result of forcings and fluctuations occurring
522 long ago (Heimbach et al., 2011), greatly complicating the interpretation of ongoing changes.

523 The results here have all been biased towards an optimistic outcome: using the estimated
524 fields both to determine the optimal linear predictors and to test them; usually retaining appar-
525 ent trends; and by employing very large scale integrals such as basin-wide transports. Consistent
526 with the earlier study of the linear predictability of the North Atlantic Oscillation (NAO; Wun-
527 sch, 1999), little skill beyond a year is found. Major elements of the ocean circulation are of
528 course, predictable far beyond that time interval: it is a very safe prediction that the thermo-
529 cline depth, the net heat content, etc. will be little changed in a decade or longer, probably
530 undetectably so, given the nature of the observing system and the natural noise.

531 In their comparison of three different model calculations of the Atlantic MOC, Bingham et
532 al. (2007) drew conclusions that are broadly similar to those found here, albeit differing in the
533 details. They found essentially no correlation in their three models between the MOC in the
534 subpolar and subtropical gyres, but did succeed in identifying a weak (relative to the overall
535 variability) lowest singular vector (EOF) representing a coupling of the two. The duration
536 of observations required to detect it was not estimated, but would clearly be extremely long
537 compared to any existing records.

538 Lack of correlation seen between the subtropical and subpolar regions can be understood
539 in rather simple terms: as discussed e.g., in Wunsch (2011) for the same state estimate, the
540 dynamical time scales for adjustment of disturbances grows very rapidly with latitude beyond
541 about 40° , so that finding simple lag correlations between gyres would be very surprising. Over
542 16 years, the subtropical gyre was found to be in near equilibrium with the wind forcing, while
543 subpolar regions were not—consistent with the time-scale growth.

544 Climate change is a global phenomenon, integrating at any given location changes originating
545 from diverse regions of the globe, not just locally, and the spatially de-correlating local responses
546 represent a summation over all times and space. If there is a timescale beyond which the MOC
547 shows large meridional coherences and/or coherence with SST as in the conveyor “ribbon”
548 cartoons, it appears to lie beyond the duration of any existing record. Sustenance over many
549 decades (Rossby et al., 2005, is an example) of globally distributed, top-to-bottom, observations
550 is urgently required, although a “hard-sell.”

551 *Acknowledgments.* Written in honor of Tom Rossby—a true pioneer of the observed-ocean.
552 Supported in part by ECCO-GODAE and the North Atlantic Meridional Overturning Circula-

553 tion Program (both NASA-funded) and our many partners in that effort. Thanks to C. King
554 and D. Spiegel for help with the model output and data sets. Comments from P. Huybers, R.
555 E. Davis, P. Heimbach, and the very detailed suggestions of two referees are appreciated.

556 **Appendix A. The Vector Least-Squares Approach to Prediction**

557 *Scalar Time Series*

558 Much of the conceptual underpinning of the standard regression methods can be simplified
 559 through a vector-least-squares point of view. The advantage is that least-squares permits a
 560 very general, and flexible, method to deal with, among other problems, the underdetermined
 561 problem of more regressors than regressees, and introduces the empirical orthogonal functions
 562 (EOFs) naturally via the singular value decomposition. Nothing that follows is original, but is
 563 a heuristic description of discrete stationary time series as discussed in innumerable textbooks.

564 Consider a stochastic zero-mean anomaly, $\xi(t = t_{now})$, where t_{now} represents the instant in
 565 time when the value of ξ is known, as are its past values, and one seeks to predict its future
 566 behavior. The time scale is chosen so that the interval is $\Delta t = 1$. Form a vector from $\xi(t)$ as

$$\boldsymbol{\xi}(t) = [\dots\xi(t-q), \xi(t-q+1), \dots, \xi(t-1), \xi(t)]^T$$

567 that is constructed from its formally infinite past and terminating at $t = t_{now}$. Let L be the actual
 568 number of observed elements, including the one at the present time. Superscript T denotes the
 569 transpose in the convention that, unless otherwise stated, all vectors have column form. Define
 570 a second vector in which everything is shifted to the right, dropping the most recent value,

$$\boldsymbol{\xi}(t-1) = [\dots\xi(t-q), \xi(t-q+1), \dots, \xi(t-2), \xi(t-1)]^T$$

571 —and which of necessity, in practice will have only $L-1$ non-zero elements. The collection of
 572 all such vectors, $\boldsymbol{\xi}(t-p) = [\dots\xi(t-q), \xi(t-q+1), \dots, \xi(t-p-1), \xi(t-p)]^T$ is a generally non-
 573 orthogonal set, noting that in observational practice the last one, $\boldsymbol{\xi}(L)$, will have only one non-
 574 zero element. This, and other, observed long-lagged vectors will thus be poor approximations
 575 to the theoretically defined semi-infinite vector.

576 Now consider another formally defined vector derived from $\xi(t)$, $\xi(t+\tau)$ as a semi-infinite
 577 one, $\tau > 0$,

$$\boldsymbol{\xi}(t+\tau) = [\dots\xi(t+\tau-q), \dots, \xi(t+\tau-1), \xi(t+\tau)]^T,$$

578 that is displaced in the opposite time direction relative to $\xi(t)$ and including the unknown *future*
 579 values

$$\xi(t+1), \xi(t+2), \dots, \xi(t+\tau).$$

$\boldsymbol{\xi}(t+\tau)$ is just another vector, and unless it is orthogonal to the collection of known past vectors,
 $\boldsymbol{\xi}(t-p)$, $p \geq 0$, one should be able to at least partially represent it in those non-orthogonal

vectors,

$$\begin{aligned} \boldsymbol{\xi}(t + \tau) = & \hspace{20em} (9) \text{ \{vector1\}} \\ & \alpha(\tau) \boldsymbol{\xi}(t) + \alpha(\tau + 1) \boldsymbol{\xi}(t - 1) + \dots + \alpha(\tau + K) \boldsymbol{\xi}(t - K) + \boldsymbol{\varepsilon}(t + \tau), \quad K \leq L - 1. \end{aligned}$$

580 The $\alpha(\tau)$ are simply the coefficients of the vector expansion, and for a stationary process would
 581 depend only upon τ , and not t . $\boldsymbol{\varepsilon}(t + \tau)$ is an error representing any elements of $\boldsymbol{\xi}(t + \tau)$ that
 582 are orthogonal to the expansion vectors (and which are the τ -lead-time “prediction error”).
 583 Determining how far back into the past, $t - K$, one should carry Eq. (9) is an important part of
 584 the inferential process. Clearly as K approaches L , the number of zero elements in the expansion
 585 vectors grows, and the particular $\boldsymbol{\xi}(t - K)$ will be a poor representation of the true vector. One
 586 prefers, $K \ll L$. Similarly, physical insight comes into the discussion, as Eq. (9) is a finite
 587 difference equation and will typically be an approximation to some partial differential system
 588 describing the time (and space) evolution of the elements $\boldsymbol{\xi}(t + \tau)$.

The simplest case is $K = 1$, and $\tau = 1$, and writing it out in full, one has,

$$\left\{ \begin{array}{c} \xi(t-1) \\ \xi(t-2) \\ \cdot \\ \cdot \\ \xi(t-(L-1)) \end{array} \right\} a_1 + \left[\begin{array}{c} \varepsilon(t) \\ \varepsilon(t-1) \\ \cdot \\ \cdot \\ \varepsilon(t-(L-2)) \end{array} \right] = \left[\begin{array}{c} \xi(t) \\ \xi(t-1) \\ \cdot \\ \cdot \\ \xi(t-(L-2)) \end{array} \right]. \quad (10) \text{ \{vector2\}}$$

$$\text{or, } \mathbf{E}_1 \mathbf{x} + \boldsymbol{\varepsilon} = \mathbf{d}, \quad \mathbf{x} = a_1, \quad (11)$$

589 The maximum number of equations is $L-1$, involving the past data as far back as $\xi(t - (L - 1))$.

An alternative formulation is Eq. (3) in the text:

$$\mathbf{E}\mathbf{x} = \mathbf{y}, \quad \mathbf{E} = \left\{ \begin{array}{cccccc} \xi(t-1) & 1 & 0 & \cdot & 0 & 0 \\ \xi(t-2) & 0 & 1 & \cdot & 0 & 0 \\ \cdot & 0 & 0 & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \xi(t-(L-1)) & 0 & 0 & \cdot & 0 & 1 \end{array} \right\}, \quad (12)$$

$$\mathbf{x} = \left[\begin{array}{c} a_1 \\ \varepsilon(t) \\ \varepsilon(t-1) \\ \cdot \\ \varepsilon(t-(L-2)) \end{array} \right], \quad \mathbf{y} = \left[\begin{array}{c} \xi(t) \\ \xi(t-1) \\ \xi(t-2) \\ \cdot \\ \cdot \\ \xi(t-(L-2)) \end{array} \right],$$

590 which identifies the values of $\varepsilon(r)$ as explicit unknowns. Now $\mathbf{E} = \{\mathbf{E}_1|\mathbf{I}\}$.

The conventional least-squares solution is (e.g., Wunsch, 2006),

$$\begin{aligned}\tilde{\mathbf{x}} = \tilde{a}_1 &= (\mathbf{E}_1^T \mathbf{E}_1)^{-1} \mathbf{E}_1^T \mathbf{d} \\ &= \frac{1/L \sum_{q=0}^{L-2} \xi(t-q-1) \xi(t-q)}{1/L \sum_{q=0}^{L-2} \xi(t-q-1)^2}\end{aligned}\tag{13} \quad \{\text{ls1}\}$$

591 which minimizes $\tilde{\boldsymbol{\varepsilon}}^T \tilde{\boldsymbol{\varepsilon}}$. The one-step prediction error (PE) is, $\tilde{\boldsymbol{\varepsilon}} = \mathbf{d} - \mathbf{E}_1 \tilde{\mathbf{x}}$. The tildes are used
 592 as a reminder that the solution is an estimate. As in any other least-squares problem, one must
 593 test the residuals, $\tilde{\boldsymbol{\varepsilon}}$, for a white-noise character. If $\tilde{\boldsymbol{\varepsilon}}$ passes that test, it is described simply by its
 594 variance, $\sigma_{\tilde{\boldsymbol{\varepsilon}}}^2$. Ordinary least-squares (e.g., Lawson and Hanson, 1995; Wunsch, 2006) produces
 595 estimates of the expected error in $\tilde{\mathbf{x}}$, etc. Quantities such as $(1/L) \sum_{q=1}^{L-1} \xi(t-q+1) \xi(t-q)$
 596 in Eq. (13) are the empirical autocovariances of $\xi(t)$ and the most conventional approach to
 597 these problems (e.g., Box et al., 2008; Priestley, 1982) formulates the problem explicitly by
 598 invoking the covariances—which are the dot (inner) products of the expansion vectors in the
 599 Yule-Walker equations. To the extent that the autocovariances are not independently known
 600 e.g., from a theory, most estimation algorithms in practice resort to forms of least-squares. Note
 601 that vectors generated from white noise sequences are orthogonal. The Kolmogoroff-Wiener-
 602 Levinson-... approach is recovered by letting $L \rightarrow \infty$, that is, the theory assumes the infinite
 603 past is known, while practice copes with a finite observed past.

604 Suppose $\tilde{\boldsymbol{\varepsilon}}$ fails the white noise test. The obvious remedy would be to try using a second
 605 vector, $\xi(t-2)$, in the expansion to remove more of the structure, so that,

$$\left\{ \begin{array}{cc} \xi(t-1) & \xi(t-2) \\ \xi(t-2) & \xi(t-3) \\ \cdot & \cdot \\ \cdot & \cdot \\ \xi(t-(L-2)) & \xi(t-(L-1)) \end{array} \right\} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \cdot \\ \cdot \\ \varepsilon(t-(L-3)) \end{bmatrix} = \begin{bmatrix} \xi(t) \\ \xi(t-1) \\ \cdot \\ \cdot \\ \xi(t-(L-3)) \end{bmatrix}, \tag{14} \quad \{\text{vector4}\}$$

606 represents an AR(2) process, which in scalar form is,

$$\xi(t) = a_1 \xi(t-1) + a_2 \xi(t-2) + \varepsilon(t).\tag{15} \quad \{\text{ar2}\}$$

607 Suppose a satisfactory (acceptable) fit has been found and so that one has estimates, \tilde{a}_1, \tilde{a}_2 , and
 608 $\tilde{\sigma}_{\tilde{\boldsymbol{\varepsilon}}}^2$. Omitting the tildes, but remembering always that all parameters are estimates, one can
 609 consider the one-step ahead prediction problem. In Eq. (15) everything is known at time $t+1$
 610 except $\varepsilon(t+2)$, which has zero-mean. Thus the best prediction is,

$$\tilde{\xi}(t+1) = a_1 \xi(t) + a_2 \xi(t-1) + 0,$$

611 and whose mean square error would be $\langle \varepsilon(t+1)^2 \rangle = \sigma_\varepsilon^2$. The two-step ahead prediction would
 612 be

$$\tilde{\xi}(t+2) = a_1 \tilde{\xi}(t+1) + a_2 \xi(t) + 0$$

613 and for which the prediction error variance is $(a_1^2 + 1) \sigma_\varepsilon^2$. This process can be continued in-
 614 definitely, the prediction error variance increasing monotonically with the prediction horizon,
 615 but never exceeding the variance of $\xi(t)$ itself: the worst prediction is $\tilde{\xi}(t+\tau) = 0$ and whose
 616 expected error is the variance of ξ ,

$$\langle \xi(t') \xi(t') \rangle = R(0) = \frac{\sigma_\varepsilon^2 (1 - a_2)}{(1 - a_1^2 - a_2^2)(1 - a_2) - 2a_1^2 a_2}.$$

617 See the references. Alternatively, one can transform $\xi(t)$ into the MA form as described in the
 618 text.

619 The formal coefficient matrices \mathbf{E}_1 or \mathbf{E} involve the observed $\xi(t)$ and inevitably contain
 620 errors. Linear least-squares treats \mathbf{E} as perfectly known, but many methods are available for
 621 discussing and remedying the bias and other errors introduced by errors in E , leading to non-
 622 linear methods (e.g., Total Least Squares; Van Huffel and Vandewalle, 1991), but which are not
 623 discussed here.

624 **Appendix B An Ensemble AR(1)**

625 An artificial AR(1), $x(t+1) = 0.3x(t) + \varepsilon(t)$, was generated for 160 samples (10-times the now
 626 available record length). The corresponding MA form has, exactly, $b_j = 0.3^j$, $j = 0, 1, \dots$

627 Thus a_1 is known exactly, as is $\varepsilon(t)$ (generated using a pseudo-random Gaussian algorithm
 628 with variance of 1). The resulting record was then divided into 10 segments each of 16 samples
 629 ($i = 1$ to 16), and the resulting system solved for $\tilde{a}_1^{(i)}$ and the estimated $\varepsilon^{(i)}(t)$. The record
 630 variance is $\langle \xi(r)^2 \rangle = \sigma_\varepsilon^2 / (1 - a_1^2)$. Fig. 19 shows the results of this experiment: The values of
 631 the estimated a_i and equivalently the b_i can and do differ substantially from the known exact
 632 values and the calculated prediction error, measured either as one-time step ahead, or as the
 633 segment record variance, varies by more than a factor of 3 from one realization to the next.
 634 They do not vary by an order of magnitude, and so one might interpret any results with the
 635 real records (below) as providing an order of magnitude estimate.

636 The variability of these estimates is known from the textbook discussions to depend upon the
 637 magnitude of a_1 (and that in turn depends directly upon the lag one covariance). With $a_1 = 0.3$,
 638 only about 9% of the variance from one time step to the next is correlated. Fig. 20 shows similar
 639 results for $a_1 = 0.9$ where about 80% of the variance would be so correlated. The coefficients

640 determined from each realization are more stable, but the prediction error (PE) growth (Fig.
641 21) with time is more rapid because small errors will persist longer, and the variance of $\xi(r)$ is
642 also greater, being proportional to $1/(1 - a_1^2)$.

643 Appendix C. Vector AR and Singular Value Decomposition

644 Consider now a generalization whereby $\xi_{i_0}(t)$ is e.g., the MOC at latitude i_0 , and one makes
645 the plausible assumption that it is correlated with, and hence predictable from, its L present
646 and past values at several other latitudes, $j = 1$ to J (including i_0). As an example, consider a
647 vector AR(2), using only two latitudes, i_0 and j , and one can write e.g.,

$$\xi_{i_0}(t) = a_1 \xi_{i_0}(t-1) + a_2 \xi_{i_0}(t-2) + b_1 \xi_j(t-1) + b_2 \xi_j(t-2) + \dots + \varepsilon(t),$$

or in matrix-vector form,

$$\begin{bmatrix} \xi_{i_0}(t) \\ \xi_{i_0}(t-1) \\ \xi_{i_0}(t-2) \\ \vdots \\ \xi_{i_0}(t-(L-3)) \end{bmatrix} = \tag{16} \text{\{vectorar1\}}$$

$$\begin{bmatrix} \xi_{i_0}(t-1) & \xi_{i_0}(t-2) & \xi_j(t-1) & \xi_j(t-2) \\ \xi_{i_0}(t-2) & \xi_{i_0}(t-3) & \xi_j(t-2) & \xi_j(t-3) \\ \xi_{i_0}(t-3) & \xi_{i_0}(t-4) & \xi_j(t-3) & \xi_j(t-4) \\ \vdots & \vdots & \vdots & \vdots \\ \xi_{i_0}(t-(L-2)) & \xi_{i_0}(t-(L-1)) & \xi_j(t-(L-2)) & \xi_j(t-(L-1)) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{bmatrix} + \begin{bmatrix} \varepsilon(t) \\ \varepsilon(t-1) \\ \varepsilon(t-3) \\ \vdots \\ \varepsilon(t-(L-3)) \end{bmatrix}, \tag{17}$$

648 where j is any other MOC time series at any latitude (or any other measured variable anywhere).
649 The b_i should not be confused with the MA coefficients used in the text. If the vector AR is
650 order N , and there are J measured time series (including the one being predicted), the equation
651 set (16) has $L - 1$ equations in $J(L - 1)$ formal unknowns (not counting the $\varepsilon(r)$). Thus
652 an AR(1) using all 111 latitudes at one degree spacing between 30°S and 80°N of estimated
653 MOC would have 111 unknowns in each of the 16 annual mean observations, leaving it greatly
654 underdetermined, with an increasing number of unknowns with any higher order AR.

655 The singular value decomposition (SVD) can be used to solve such underdetermined problems
656 (e.g., Wunsch, 2006). The coefficient matrix made up of the expansion vectors (the regressors),

657 is written in canonical form as

$$\mathbf{E} = \lambda_1 \mathbf{u}_1 \mathbf{v}_1^T + \lambda_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \lambda_K \mathbf{u}_K \mathbf{v}_K^T, \quad (18) \quad \{\text{svd2}\}$$

658 where the $\mathbf{u}_i, \mathbf{v}_i$ are the orthonormal singular vectors, and the λ_i are the singular values. $K \leq 15$,
659 is the maximum possible rank of \mathbf{E} here. The \mathbf{u}_i are often known as empirical orthogonal
660 functions (EOFs) and corresponding \mathbf{v}_i are the temporal coefficients. λ_i^2 is the contribution to
661 the squared norm of \mathbf{E} .

662 In the present case, the singular value decomposition shows that \mathbf{E} is formally of full rank,
663 $K = 15$, and at full rank, \mathbf{E} is exactly represented by 15 pairs of orthonormal vectors in Eq. (18).
664 A more plausible estimate of the useful rank is either 9 or 13, depending upon how large the
665 noise is estimated to be. $K = 9$ suggests approximately nine independent pieces of information
666 amongst the 111 latitudinal values of the MOC at a one-year time lag. The SVD solution is,

$$\tilde{\mathbf{x}} = \mathbf{v}_1 (\mathbf{u}_1^T \mathbf{y} / \lambda_1) + \mathbf{v}_2 (\mathbf{u}_2^T \mathbf{y} / \lambda_2) + \dots + \mathbf{v}_K (\mathbf{u}_K^T \mathbf{y} / \lambda_K), \quad (19) \quad \{\text{svdsol1}\}$$

667 but results from this approach are not shown here, as they founder on the same too-short record
668 duration.

References

- 670 Anderson, D.L.T., Bryan, K., Gill, A.E., Pacanowski, R.C., 1979. Transient-response of the
671 North Atlantic - some model studies. *Journal of Geophysical Research-Oceans and Atmospheres*
672 84, 4795-4815.
- 673 Bengtsson, L., Hagemann, S., Hodges, K.I., 2004. Can climate trends be calculated from re-
674 analysis data? *Journal of Geophysical Research-Atmospheres* 109, DOI 10.1029/2004jd004536.
- 675 Bingham, R.J., Hughes, C.W., Roussenov, V., Williams, R.G., 2007. Meridional coherence of
676 the North Atlantic meridional overturning circulation. *Geophysical Research Letters* 34, L23606
677 DOI10.1029/2007gl031731.
- 678 Box, G.E.P., Jenkins, G.M., Reinsel, G.C., 2008. *Time Series Analysis : Forecasting And Con-*
679 *trol*, 4th ed. John Wiley, Hoboken, N.J.
- 680
- 681 Branstator, G., Teng, H.Y., 2010. Two Limits of Initial-Value Decadal Predictability in a
682 CGCM. *Journal of Climate* 23, 6292-6311.
- 683 Bromwich, D.H., Fogt, R.L., Hodges, K.I., Walsh, J.E., 2007. A tropospheric assessment of
684 the ERA-40, NCEP, and JRA-25 global reanalyses in the polar regions. *Journal of Geophysical*
685 *Research-Atmospheres* 112, DOI 10.1029/2006jd007859.
- 686 Davis, R.E., 1976. Predictability of sea-surface temperature and sea-level pressure anomalies
687 over North Pacific Ocean. *Journal of Physical Oceanography* 6, 249-266.
- 688 Davis, R.E., 1978. Predictability of sea-level pressure anomalies over North Pacific Ocean.
689 *Journal of Physical Oceanography* 8, 233-246.
- 690 Davis, R.E., 1979. Search for short-range climate predictability. *Dynamics of Atmospheres and*
691 *Oceans* 3, 485-497.
- 692 Deser, C., Alexander, M.A., Xie, S.P., Phillips, A.S., 2010. Sea surface temperature variability:
693 patterns and mechanisms. *Annual Review of Marine Science* 2, 115-143.
- 694 Gebbie, G. and Huybers. P., 2011. How is the ocean filled? *Geophysical Research Letters*, 38,
695 DOI10.1029/2011GL046769.
- 696 Hamilton, J.D., 1994. *Time Series Analysis*. Princeton Un. Press.
- 697 Hannan, E.J., 1970. *Multiple Time Series*. John Wiley, New York.
- 698 Heimbach, P., Wunsch, R. M. Ponte, G. Forget, C. Hill and J. Utke, 2011. Timescales and
699 regions of the sensitivity of Atlantic meridional volume and heat transport magnitudes: toward
700 observing system design, *Deep-Sea Research-II* 58, 1858-1879 2011.
- 701 Hurrell J. and Co-authors, 2010. Decadal climate prediciton: opportunities and challenges,
702 in *Proceedings of Ocean Obs09: Sustained Ocean Observations and Information for Soci-*

703 ety, Vol. 2, Hall, J., Harrison, D. E. and D. Stammer, Eds., ESA Publication WPP-306,
704 doi:10.5270/OceanObs09.cwp.45

705 Kanzow, T., Johnson, H.L., Marshall, D.P., Cunningham, S.A., Hirschi, J.J.M., Mujahid, A.,
706 Bryden, H.L., Johns, W.E., 2009. Basinwide integrated volume transports in an eddy-filled
707 ocean. *Journal of Physical Oceanography* 39, 3091-3110.

708 Lawson, C.L., Hanson, R.J., 1995. *Solving Least Squares Problems*. SIAM, Philadelphia.

709 Ljung, L., 1999. *System Identification: Theory for the User*, 2nd Ed. Prentice-Hall, Englewood
710 Cliffs, N. J.

711 Lorbacher, K., Dengg, J., Boning, C.W., Biastoch, A., 2010. Regional patterns of sea level
712 change related to interannual variability and multidecadal trends in the Atlantic meridional
713 overturning circulation. *Journal of Climate* 23, 4243-4254.

714

715 Meehl, G.A., Goddard, L., Murphy, J., and 17 others., 2009. Decadal prediction: Can it be
716 skillful? *Bulletin of the American Meteorological Society* 90, 1467-+.

717 Msadek, R., Dixon, K.W., Delworth, T.L., Hurlin, W., 2010. Assessing the predictability of the
718 Atlantic meridional overturning circulation and associated fingerprints. *Geophysical Research*
719 *Letters* 37, L19608, DOI10.1029/2010gl044517

720 Nelles, O., 2001. *Nonlinear System Identification: From Classical Approaches to Neural Net-*
721 *works and Fuzzy Models*. Springer, Berlin.

722 Newman, M., 2007. Interannual to Decadal Predictability of Tropical and North Pacific Sea
723 Surface Temperatures. *Journal of Climate* 20, 2333-2356.

724 Priestley, M.B., 1982. *Spectral Analysis and Time Series*. Volume 1: Univariate Series. Volume
725 2: Multivariate Series, Prediction and Control. Academic, London.

726 Reynolds, R.W., Smith, T.M., 1995. A high-resolution global sea-surface temperature climatol-
727 ogy. *Journal of Climate* 8, 1571-1583.

728 Robinson, E.A., 1981. *Time Series Analysis And Applications*. Goose Pond Press, Houston,
729 Tex.

730 Rossby, T., Flagg, C.N., Donohue, K., 2005. Interannual variations in upper-ocean transport
731 by the Gulf Stream and adjacent waters between New Jersey and Bermuda. *Journal of Marine*
732 *Research* 63, 203-226.

733 Rossby, T., Flagg, C.N., Donohue, K., 2010. On the variability of Gulf Stream transport from
734 seasonal to decadal timescales. *Journal of Marine Research*, 68, 503-522.

735 Storch, H.v., Zwiers, F.W., 2001. *Statistical Analysis In Climate Research*, 1st pbk. ed. Cam-
736 bridge University Press.

737

738 Tizard, T.H., Mosely, H.N., Buchanan, J.Y., Murray, J., 1885. Report on the Scientific Results
739 of the Voyage of H. M. S. Challenger During the Years 1873-1876, Narrative of the Cruise of H.
740 M. S. Challenger with a General Account of the Scientific Results of the Expedition, 1, First
741 Part. H. M. Stationery Office, Edinburgh.

742 Tourre, Y.M., Rajagopalan, B., Kushnir, Y., 1999. Dominant patterns of climate variability in
743 the Atlantic Ocean during the last 136 years. *Journal of Climate* 12, 2285-2299.

744 Van Huffel, S., Vandewalle, J., 1991. The Total Least Squares Problem. *Computational Aspects*
745 *and Analysis*. SIAM, Philadelphia.

746 Veronis, G., Stommel, H., 1956. The action of variable wind stresses on a stratified ocean.
747 *Journal of Marine Research*, 15, 43-75.

748

749 Vinogradova, N.T., Ponte, R.M., and Heimbach, P., 2010: Dynamics and forcing of sea surface
750 temperature variability on climate time scales. *Journal of Climate*, submitted.

751 Wang, W.Q., Kohl, A., Stammer, D., 2010. Estimates of global ocean volume transports during
752 1960 through 2001. *Geophysical Research Letters* 37.

753 Woollings, T., Hoskins, B., Blackburn, M., Hassell, D., Hodges, K., 2010. Storm track sensitivity
754 to sea surface temperature resolution in a regional atmosphere model. *Climate Dynamics* 35,
755 341-353.

756 Wunsch, C., 1999. The interpretation of short climate records, with comments on the North
757 Atlantic and Southern Oscillations. *Bulletin of the American Meteorological Society* 80, 245-
758 255.

759 Wunsch, C., 2006. *Discrete Inverse And State Estimation Problems: With Geophysical Fluid*
760 *Applications*. Cambridge University Press.

761 Wunsch, C., 2008. Mass and volume transport variability in an eddy-filled ocean. *Nature*
762 *Geoscience* 1, 165-168.

763 Wunsch, C., Heimbach, P., 2006. Estimated decadal changes in the North Atlantic meridional
764 overturning circulation and heat flux 1993-2004. *Journal of Physical Oceanography* 36, 2012-
765 2024.

766 Wunsch, C., Heimbach, P., 2007. Practical global oceanic state estimation. *Physica D-Nonlinear*
767 *Phenomena* 230, 197-208.

768 Wunsch, C., Heimbach, P., 2009. The global zonally integrated ocean circulation, 1992-2006:
769 seasonal and decadal variability. *Journal of Physical Oceanography* 39, 351-368.

770 Wunsch, C., Heimbach, P., Ponte, R.M., Fukumori, I., Members, E.-G.C., 2009. The global
771 general circulation of the ocean estimated by the ECCO-consortium. *Oceanography* 22, 88-103.

772 Yin, J.J., Schlesinger, M.E., Stouffer, R.J., 2009. Model projections of rapid sea-level rise on

773 the northeast coast of the United States. *Nature Geoscience* 2, 262-266.
774 Zanna, L., Heimbach, P., Moore, A.M., Tziperman, E., 2011. Optimal excitation of interannual
775 Atlantic meridional overturning circulation variability. *Journal of Climate* 24, 413-427.
776 Zhang, H.H., Wu, L.X., 2010. Predicting North Atlantic sea surface temperature variability
777 on the basis of the first-mode baroclinic Rossby wave model. *Journal of Geophysical Research-*
778 *Oceans* 115, C09030, DOI10.1029/2009jc006017.

779

Figure and Table Captions

781 Table 1. Summary statistics. Variances are either in $^{\circ}\text{C}^2$ (for SST) or Sv^2 for the meridional
 782 overturning circulation (MOC). PE is the prediction error. The record variance is not the sum of
 783 the component variances because the monthly values include the low frequency variability. Some
 784 prediction error values are omitted as being of no particular interest. GBB denotes the Grand
 785 Bahama Bank square, and ECCO is the consortium Estimating the Circulation and Climate of
 786 the Ocean. MA(M) indicates that the prediction error was deduced by converting the AR(1)
 787 model into an MA of order M.}

788 Fig. 1. Sixteen year time mean sea surface temperature (SST, in $^{\circ}\text{C}$) from the ECCO-
 789 GODAE estimate in the North Atlantic. Small white square, called the Grand Banks Box—
 790 GBB, is used as prototypical of the areal prediction problem.

791 2. The Grand Banks Box (GBB) area average temperature, $T_G(t)$ (solid curve), the best-
 792 fitting annual cycle including its first three harmonics (dashed), and the monthly residuals of
 793 the annual cycle (dotted). Start is 1992.

794 3. (a) Periodogram of $T_G(t)$ for the ECCO estimate (dashed) and longer Reynolds and Smith
 795 (1995) time series (solid curve). (b) Cumulative integral of the periodgrams in (a) normalized
 796 to a sum of 1, so that the dominance by the annual peak in both cases is clear. (c) Spectral
 797 estimates for both time series after removal of the annual cycle and its harmonics. The annual
 798 peak is so narrow as to be indistinguishable at this resolution from a pure sinusoid. At low
 799 frequencies, a power law of frequency to the power -2.5 is approximately correct.

800 4. Monthly values of $T'_{GBB}(t)$, (start is 1992) residual of the area average GBB SST, after
 801 removal of the annual cycle and its harmonics. The visual trend, if secular—meaning extending
 802 far beyond the record length—contributes to the apparent predictability as it is here treated as
 803 part of a red noise process. (Repeated from Fig. 2.)

804 5. Prediction error out to 6 months for $T'_{GBB}(t)$. Note that the variance of the monthly
 805 means of $T'_{GBB}(t)$ is 0.7°C^2 , which is the maximum prediction

806 6. Annual mean values for the Reynolds and Smith (1995)—solid line and the ECCO results
 807 in the GBB (dashed).

808 7. Prediction error growth in *years* for $T'_{GBB}(t)$ from an AR(1) converted to an MA(5).
 809 Total variance is 0.78°C .

810 8. Zonal and time mean meridional transport (not the stream function). The upper 200m
811 has a particularly complex structure at low latitudes (see Wunsch and Heimbach, 2009).

812 9. The maximum meridional transport, integrated from the sea surface, averaged over 16
813 years (upper panel). Lower panel shows the depth where the time-mean value is obtained.

814 10. Monthly values of $V(y, z, t)$ in Sverdrups ($10^6 \text{m}^3/\text{s}$) for a succession of Januarys showing
815 the typical interannual variability occurring at depth. Year 2 is 1993. The origin of these small
816 meridional scale features is not explored here, but may be associated with response to the Ekman
817 forcing in shallow water areas (e.g., Davis, 2010).

818 11. The monthly MOC *anomaly* without the annual mean cycle at 25°N and 50°N (a) and
819 at 20°S (b). Little visual similarity is apparent.

820 12. MOC variances (solid curve), the annual contribution (with harmonics) as a function of
821 latitude in Sv^2 (dashed line), and the residual after removal of the annual cycle (dotted).

822 13. Power density spectral estimate of monthly MOC values at three latitudes. The an-
823 nual cycle and its harmonics are visible, as is the low frequency asymptote toward white noise
824 behavior. This spectral density is, overall, nearly flat. The annual peak is broadened by the
825 multi-tapers used to form the estimated spectrum and the very lowest frequency estimate has a
826 known negative bias.

827 14. Correlation matrix with latitude of the annual mean MOC (left panel). Right panel
828 is an expanded color scale version of the left panel, showing only the apparently statistically
829 significant values. No negative correlations are significant.

830 15. Coherence amplitude and phase between 20.5°N and 50.5°N . Significant coherence van-
831 ishes at periods longer than one year. High frequency coherence is in large part that of the
832 annual cycle and its harmonics and for which the level-of-no-significance shown is inappropriate.

833 16. Coherence between the monthly MOC at 20.5°S and 50.5°N . Apart from the annual
834 cycle, where the conventional statistics do not apply, there is no significant coherence.

835 17. Prediction error of the MOC as a function of year at 50°N from a univariate $\text{AR}(1)$.

836 18. Correlation coefficient between the maximum MOC through time (annual means) with
837 the GBB SST (left panel). The last row and column are the SST correlations. Omitting
838 the last row and column repeats the values in Fig. 14. Right panel shows the same results

839 but with a linear trend removed from SST, thus reducing the correlations. No values below
840 magnitude 0.5 are statistically significant. (These correlations are with the MOC defined as
841 integrated to the time-mean maximum depth. Results with the time-varying integration depth
842 are indistinguishable.)

843 19. $\tilde{a}^{(i)}$ from each 16-element segment of the record in Fig. 19 (a) panel; true value is $a = 0.3$;
844 (b) the estimated uncertainty in those values and (c) the variance in the 15 samples estimates
845 of $\varepsilon(t)$ in the segment. The correct value is 1.

846 20. Same as Fig. 19 except for $a = 0.9$.

847 21. The 10 different realizations of the estimated MA coefficients for $a = 0.9$ (left panel), and
848 the corresponding prediction error growth through time (right panel). Each line corresponds to
849 a different 16 time step realization. True values are shown as ‘o’.

Variable	GBB SST (ECCO) $^{\circ}\text{C}^2$	GBB (Reynolds & Smith) $^{\circ}\text{C}^2$	MOC at 20°S Sv^2	MOC at 25°N Sv^2	MOC at 50°N Sv^2
Total Record	10.2	9.9	6.5	10.2	10.5
Annual cycle	9.45 (93%)	8.8	2.1	2.9	3.6
Record w/o annual cycle	0.78	0.7 (7%)	4.4	7.2	7.0
Annual averages	0.50 (5% of the total)	0.36(3.6%)	1.9	2.2	1.8
One month PE	0.2	0.3 (MA(3) and MA(10))	2.4 MA(4)	6.5 MA(4)	5.6 MA(4)
Six month PE	0.6	0.7	-	-	-
One year PE	0.2 (AR(1) with trend)	0.05 (MA(4))	0.5 MA(4)	0.2 MA(4)	0.8 MA(4)
Three year PE	0.4	0.3	-	-	1.5

Table 1: Summary statistics. Variances are either in $^{\circ}\text{C}^2$ (for SST) or Sv^2 for the meridional overturning circulation (MOC). PE is the prediction error. The record variance is not the sum of the component variances because the monthly values include the low frequency variability. Some prediction error values are omitted as being of no particular interest. GBB denotes the Grand Bahama Bank square, and ECCO is the consortium Estimating the Circulation and Climate of the Ocean. MA(M) indicates that the prediction error was deduced by converting the AR(1) model into an MA of order M.

{TableKey}

{table}

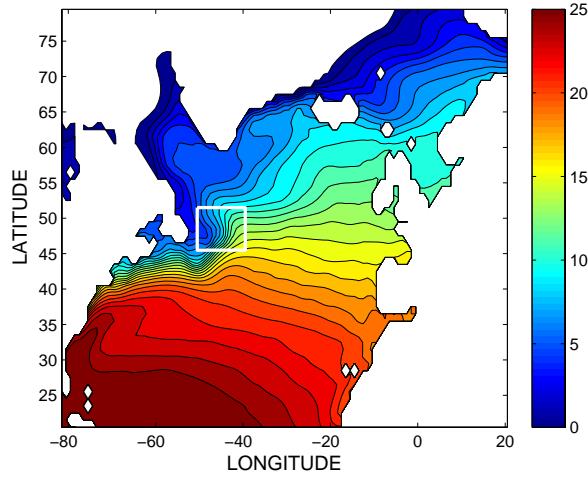


Figure 1: Sixteen year time mean sea surface temperature (SST, in $^{\circ}\text{C}$) from the ECCO-GODAE estimate in the North Atlantic. Small white square, called the Grand Banks Box—GBB, is used as prototypical of the areal prediction problem.

{sst_time_mean

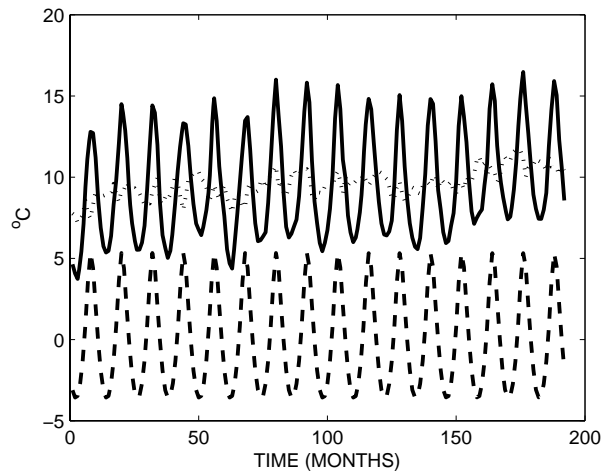


Figure 2: The Grand Banks Box (GBB) area average temperature $T_G(t)$ (solid curve), the best-fitting annual cycle including its first three harmonics (dashed), and the monthly residuals of the annual cycle (dotted). Start is 1992.

{sst_areaavg&r

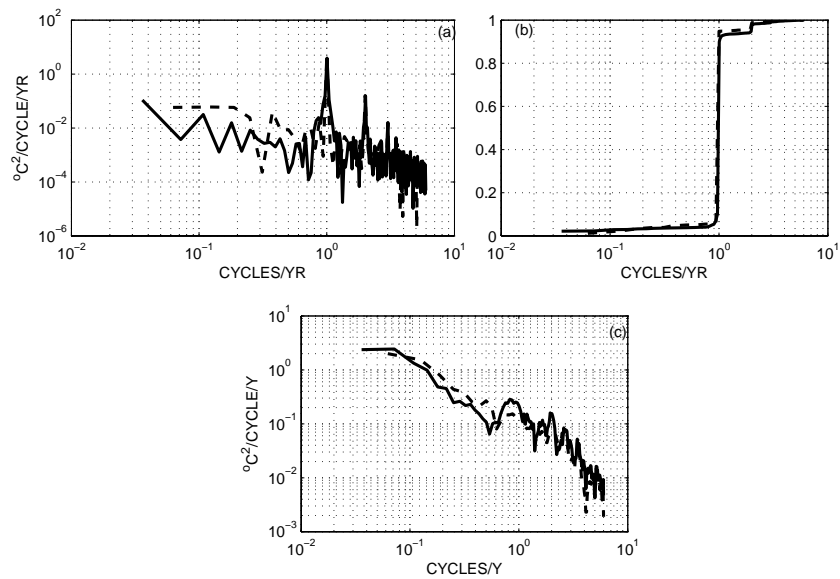


Figure 3: (a) Periodogram of $T_G(t)$ for the ECCO estimate (dashed) and longer Reynolds and Smith (1995) time series (solid curve). (b) Cumulative integral of the periodograms in (a) normalized to a sum of 1, so that the dominance by the annual peak in both cases is clear. (c) Spectral estimates for both time series after removal of the annual cycle and its harmonics. The annual peak is so narrow as to be indistinguishable at this resolution from a pure sinusoid. At low frequencies, a power law of frequency to the power -2.5 is approximately correct.

{area_periodo_

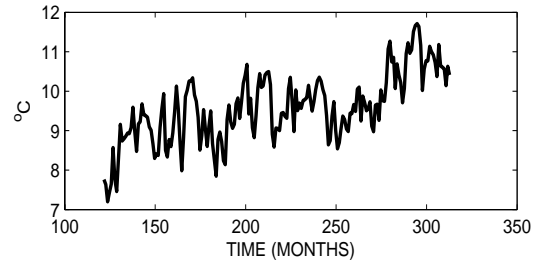


Figure 4: Monthly values of $T'_{GBB}(t)$, (start is 1992) residual of the area average GBB SST, after removal of the annual cycle and its harmonics. The visual trend, if secular—meaning extending far beyond the record length—contributes to the apparent predictability as it is here treated as part of a red noise process. (Repeated from Fig. 2.)

{sst_area_noan

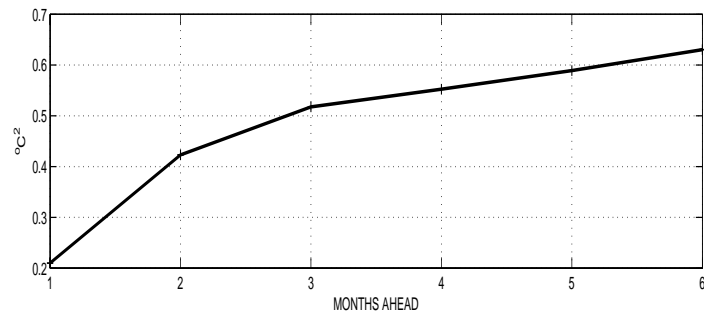


Figure 5: Prediction error out to 6 months for $T'_{GBB}(t)$. Note that the variance of the monthly means of $T'_{GBB}(t)$ is 0.7°C^2 , which is the maximum prediction error.

{gbb_pe_ma6.ep

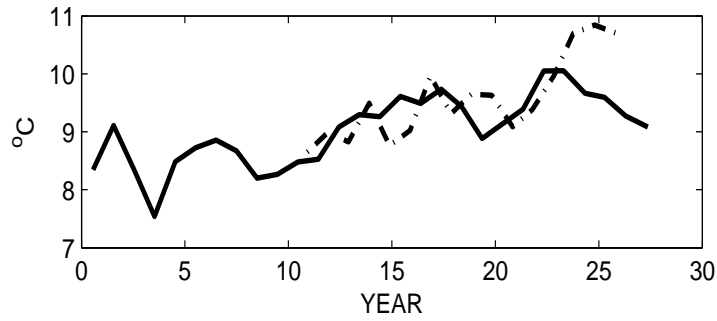


Figure 6: Annual temperature in the GBB (dashed)

ECCO results in

{area_anmeans

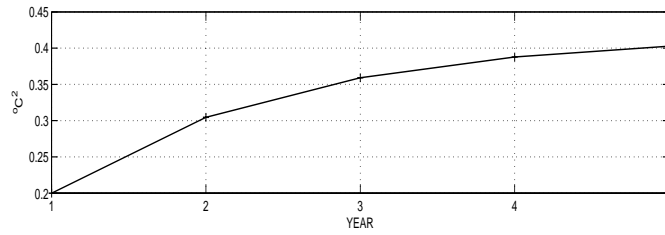


Figure 7: Prediction error growth in years for $T'_{GBB}(t)$ from an AR(1) converted to an MA(5). Total variance is 0.78°C .

{pred_error_ar

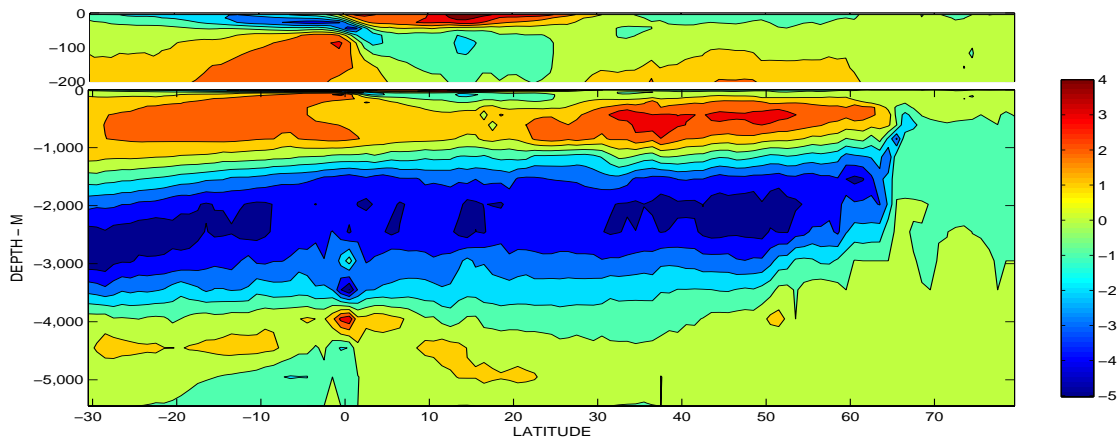


Figure 8: Zonal and time mean meridional transport (not the stream function). The upper 200m has a particularly complex structure at low latitudes (see Wunsch and Heimbach, 2009).

{moc_timemean_

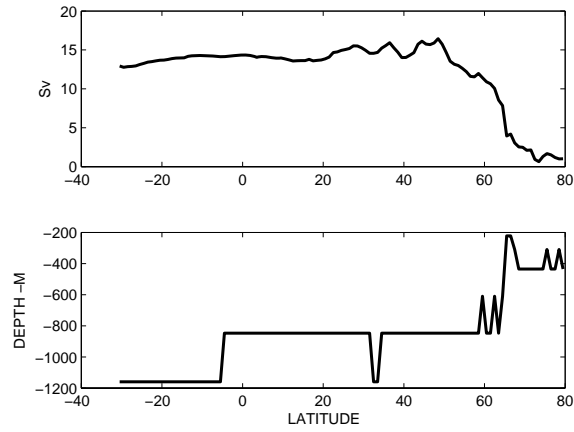


Figure 9: The maximum meridional transport, integrated from the sea surface, averaged over 16 years (upper panel). Lower panel shows the depth where the time-mean value is obtained.

{mocmax&deptht

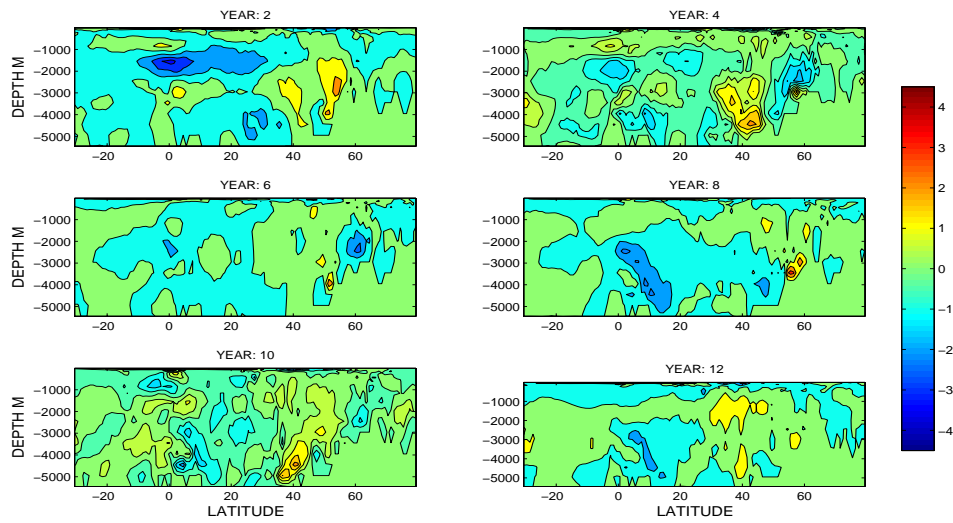


Figure 10: Monthly values of $V(y, z, t)$ in Sverdrups ($10^6 \text{m}^3/\text{s}$) for a succession of Januarys showing the typical interannual variability occurring at depth. Year 2 is 1993. The origin of these small meridional scale features is not explored here, but may be associated with response to the Ekman forcing in shallow water areas (e.g., Davis, 2010).

{moc_every2yea

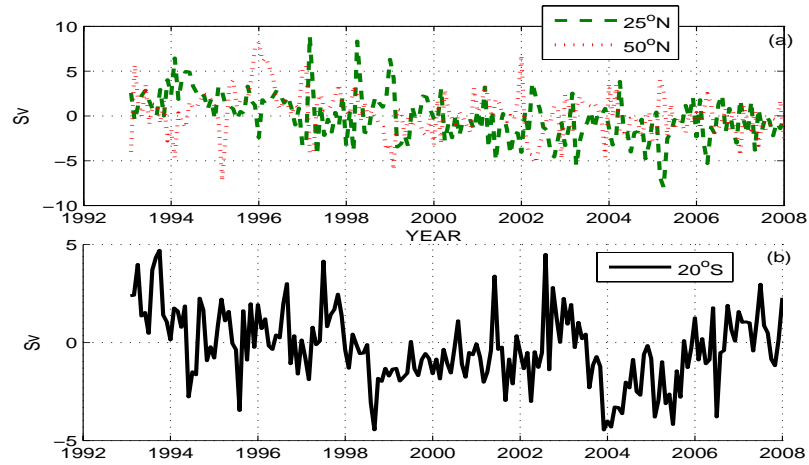


Figure 11: The maximum monthly MOC *anomaly* without the annual mean cycle at 25°N and 50°N (a) and at 20°S (b). Little visual similarity is apparent.

{moc_3lats_ts_

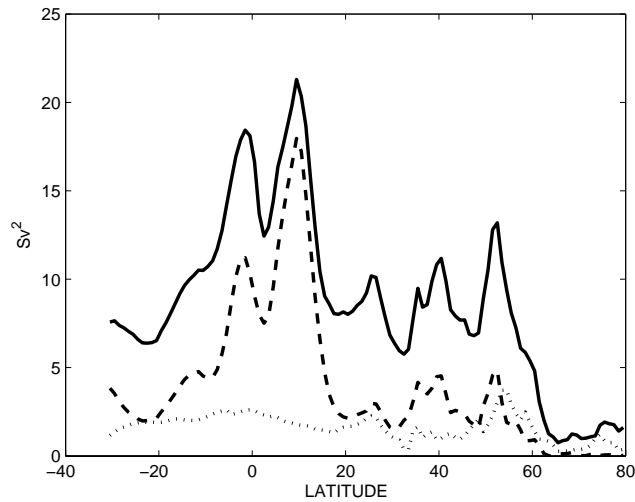


Figure 12: MOC variances (solid curve), the annual contribution (with harmonics) as a function of latitude in Sv² (dashed line), and the residual after removal of the annual cycle (dotted).

{var_all_lats&

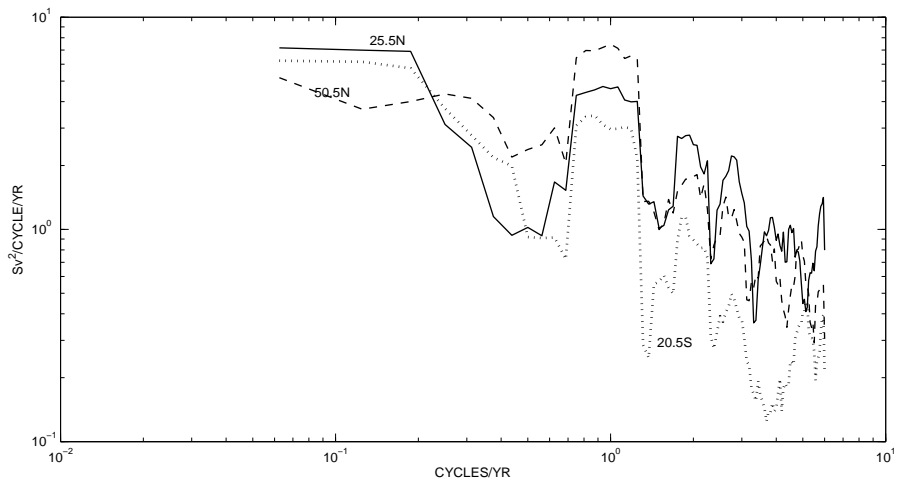


Figure 13: Power density spectral estimate of monthly MOC values at three latitudes. The annual cycle and its harmonics are visible, as is the low frequency asymptote toward white noise behavior. This spectral density is, overall, nearly flat. The annual peak is broadened by the multi-tapers used to form the estimated spectrum and the very lowest frequency estimate has a known negative bias.

{moc_pd_3lats.

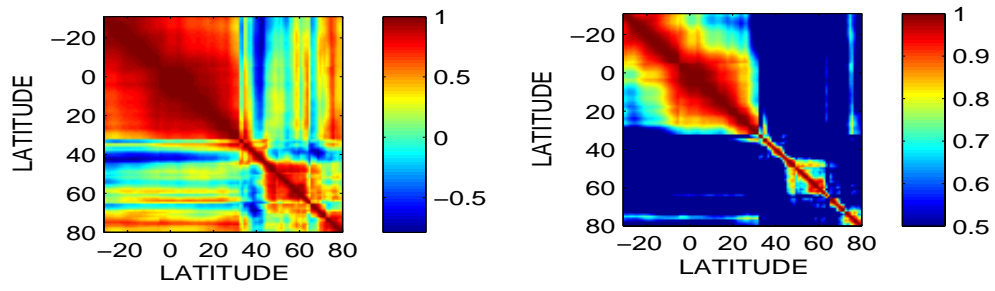


Figure 14: Correlation matrix with latitude of the annual mean MOC (left panel). Right panel is an expanded color scale version of the left panel, showing only the apparently statistically significant values. No negative correlations are significant.

{moc_latcorr_a

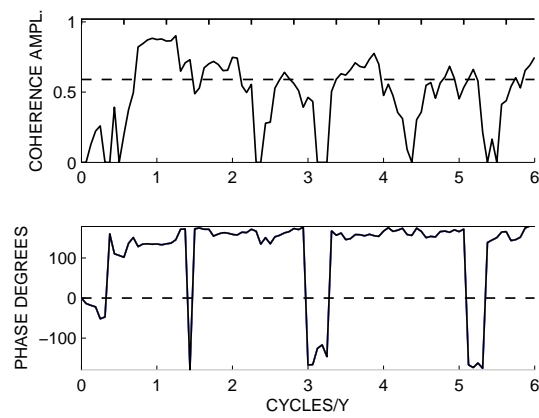


Figure 15: Coherence amplitude and phase between 20.5°N and 50.5°N . Significant coherence vanishes at periods longer than one year. High frequency coherence is in large part that of the annual cycle and its harmonics and for which the level-of-no-significance shown is inappropriate.

{moc_coher25n5

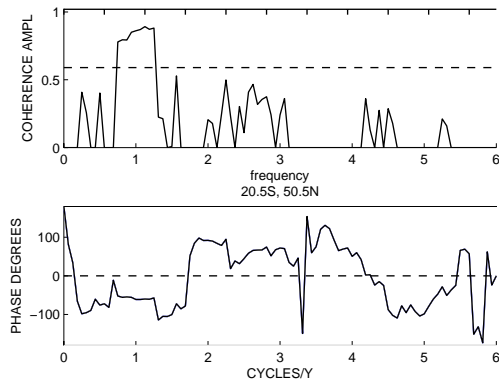


Figure 16: Coherence between the monthly MOC at 20.5°S and 50.5°N. Apart from the annual cycle, where the conventional statistics do not apply, there is no significant coherence.

{moc_coher21s5}

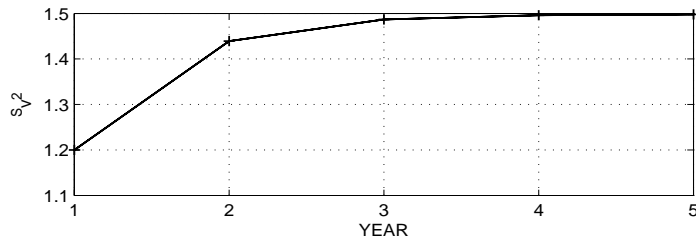


Figure 17: Prediction error as a function of year at 50°N from a univariate AR(1).

{pemoc50nfroma}

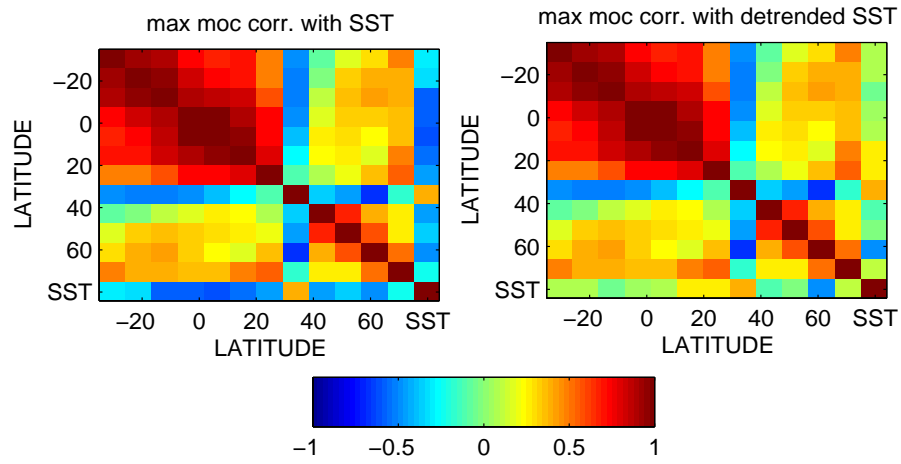


Figure 18: Correlation coefficient between the maximum MOC through time (annual means) with the GBB SST (left panel). The last row and column are the SST correlations. Omitting the last row and column repeats the values in Fig. 14. Right panel shows the same results but with a linear trend removed from SST, thus reducing the correlations. No values below magnitude 0.5 are statistically significant. (These correlations are with the MOC defined as integrated to the time-mean maximum depth. Results with the time-varying integration depth are indistinguishable.)

{mocmax_latcor

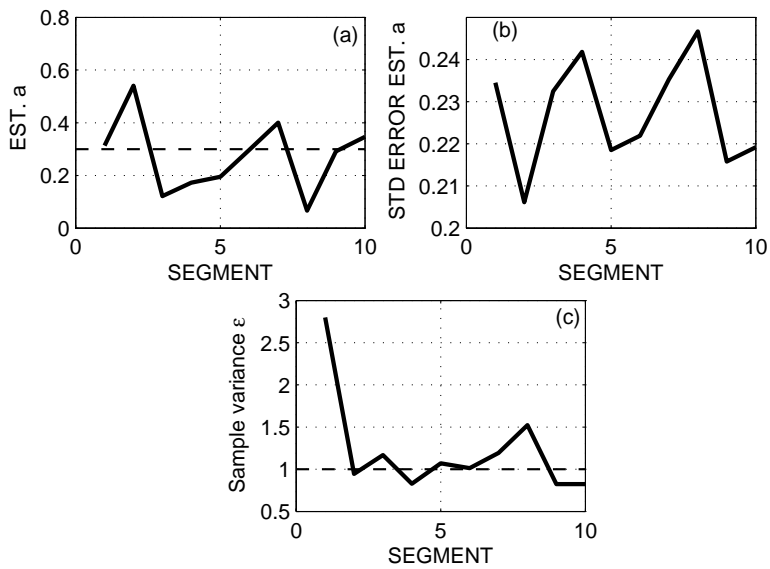


Figure 19: $\tilde{a}^{(i)}$ from each 16-element segment of the record in Fig. 19 (a) panel; true value is $a = 0.3$; (b) the estimated uncertainty in those values and (c) the variance in the 15 samples estimates of $\epsilon(t)$ in the segment. The correct value is 1.

{atru3sols.ep

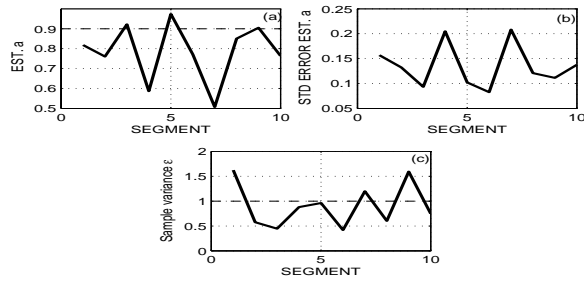


Figure 20: Same as Fig. 19 except for $a = 0.9$.

{atru9sols.ep

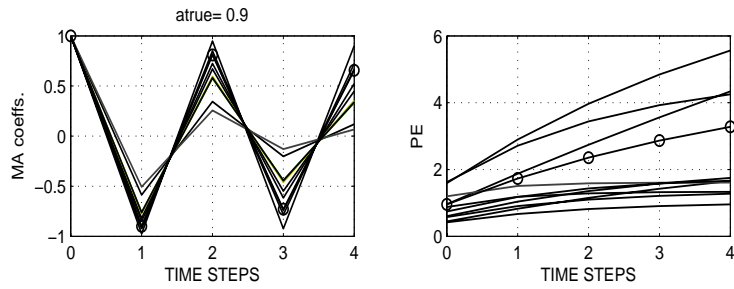


Figure 21: The 10 different realizations of the estimated MA coefficients for $a = 0.9$ (left panel), and the corresponding prediction error growth through time (right panel). Each line corresponds to a different 16 time step realization. True values are shown as 'o'.

{atru9ma&pe.e