Chapter 8

100 Years of Progress in Understanding the Dynamics of Coupled Atmosphere–Ocean Variability

DAVID S. BATTISTI

Department of Atmospheric Sciences, University of Washington, Seattle, Washington

DANIEL J. VIMONT

Atmospheric and Oceanic Science, University of Wisconsin–Madison, Madison, Wisconsin

BENJAMIN P. KIRTMAN

Department of Atmospheric Science, Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida

ABSTRACT

In situ observation networks and reanalyses products of the state of the atmosphere and upper ocean show well-defined, large-scale patterns of coupled climate variability on time scales ranging from seasons to several decades. We summarize these phenomena and their physics, which have been revealed by analysis of observations, by experimentation with uncoupled and coupled atmosphere and ocean models with a hierarchy of complexity, and by theoretical developments. We start with a discussion of the seasonal cycle in the equatorial tropical Pacific and Atlantic Oceans, which are clearly affected by coupling between the atmosphere and the ocean. We then discuss the tropical phenomena that only exist because of the coupling between the atmosphere and the ocean: the Pacific and Atlantic meridional modes, the El Niño–Southern Oscillation (ENSO) in the Pacific, and a phenomenon analogous to ENSO in the Atlantic. For ENSO, we further discuss the sources of irregularity and asymmetry between warm and cold phases of ENSO, and the response of ENSO to forcing. Fundamental to variability on all time scales in the midlatitudes of the Northern Hemisphere are preferred patterns of uncoupled atmospheric variability that exist independent of any changes in the state of the ocean, land, or distribution of sea ice. These patterns include the North Atlantic Oscillation (NAO), the North Pacific Oscillation (NPO), and the Pacific–North American (PNA) pattern; they are most active in wintertime, with a temporal spectrum that is nearly white. Stochastic variability in the NPO, PNA, and NAO force the ocean on days to interannual time scales by way of turbulent heat exchange and Ekman transport, and on decadal and longer time scales by way of wind stress forcing. The PNA is partially responsible for the Pacific decadal oscillation; the NAO is responsible for an analogous phenomenon in the North Atlantic subpolar gyre. In models, stochastic forcing by the NAO also gives rise to variability in the strength of the Atlantic meridional overturning circulation (AMOC) that is partially responsible for multidecadal anomalies in the North Atlantic climate known as the Atlantic multidecadal oscillation (AMO); observations do not yet exist to adequately determine the physics of the AMO. We review the progress that has been made in the past 50 years in understanding each of these phenomena and the implications for short-term (seasonal-to-interannual) climate forecasts. We end with a brief discussion of advances of things that are on the horizon, under the rug, and over the rainbow.

Corresponding author: David S. Battisti, battisti@washington.edu

DOI: 10.1175/AMSMONOGRAPHS-D-18-0025.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
1. Introduction

There was little discussion of coupled atmosphere–ocean variability in the first two-thirds of the twentieth century. In an early study, Sir Gilbert Walker analyzed station data around the world and coined the term “Southern Oscillation” for a large-scale coherent oscillation in sea level pressure and in precipitation in the Maritime Continent (Walker 1924).1 Decades later, Berlage (1966) linked the Southern Oscillation to episodic, localized warming off Peru and Ecuador: this warming was the El Niño phenomenon documented nearly 75 years earlier (Carranza 1892; Carrillo 1893; Pezet 1895, 1896). Bjerknes (1969) showed that changes in the trade winds along the equator in the Pacific were associated with the Southern Oscillation through the Walker circulation. Bjerknes presented evidence that changes in the strength of the trade winds were intimately related to the large-scale east–west sea surface temperature (SST)2 gradient across the Pacific; today this tight relationship is recognized as the Bjerknes feedback and is one of two central ingredients to the large-scale El Niño–Southern Oscillation (ENSO) phenomenon (the other is the ocean adiabatic adjustment to the changes in wind stress; see section 4). While earlier examples of this necessary two-way coupling may exist, the description by Bjerknes (1969) marks a beginning to the now widespread recognition of coupled ocean–atmosphere phenomena, as defined herein.

In the 1960s and 1970s, observational studies of the tropical Pacific atmosphere and ocean gave rise to a “canonical” view of ENSO as a pan-Pacific phenomenon, with warm El Niño events lasting 12–18 months or so, usually followed by cold La Niña events of lesser amplitude and lasting a few years (Rasmusson and Carpenter 1982). El Niño events occurred every 4–7 years and tended to peak at the end of the calendar year. Also in the 1960s and 1970s, theoretical advances led to an understanding of the response of the tropical oceans to wind stress forcing, and to the response of the atmosphere to changes in SST (see section 4b).

The 1982/83 El Niño event (the warm phase of ENSO) was remarkable for its amplitude and duration. It inspired meteorologists and oceanographers to come together and plan the 10-yr program Tropical Oceans on the Global Atmosphere (TOGA) to study the impact of the oceans on the atmosphere over 1985–94. TOGA significantly enhanced the observing system in the tropical Pacific [for an overview, see McPhaden et al. (1998)] and ushered in intermediate complexity models of the coupled atmosphere–ocean system that led to an understanding of the essential aspects of the canonical ENSO cycle, including the spatial structure and amplitude of the warm and cold (El Niño and La Niña) events, the period between warm events, and the seasonality in the variance of ENSO. In turn, analyses of these models led to the view that ENSO is an intrinsic mode of the dynamically coupled atmosphere–ocean system in the tropical Pacific, in which tropical ocean dynamics are essential for the evolution of SST through a rich mixture of ocean dynamics and surface fluxes. Following on the TOGA program and the World Ocean Climate Experiment (WOCE), the World Climate Research Programme outlined a science plan to advance understanding and predictability of climate variability arising from coupled ocean–atmosphere interactions (U.S. CLIVAR Scientific Steering Committee 2013). The Climate Variability and Predictability Program (CLIVAR) program (both International CLIVAR in 1995 and U.S. CLIVAR in 1997) continued to advance understanding of coupled ocean–atmosphere variations with a more global focus, including coupled ocean–atmosphere phenomena in the tropical Atlantic and Indian Oceans and in the mid- and high latitudes.

Today, a quarter century after TOGA, further observations and modeling studies have confirmed this view of ENSO (but important questions remain; see section 8). In addition, other modes of climate variability have been identified in observations and simulated by climate models that likely exist only because of a coupling between the atmosphere and ocean. For example, the seasonal cycle along the equator in the tropical Atlantic and Pacific is now understood as being born from a coupling of atmosphere and ocean in response to the seasonal cycle in insolation (section 2). The meridional modes describe patterns of intraseasonal to interannual variability in the tropical Pacific and Atlantic basins that are intrinsically forced by (stochastic) atmosphere variability, but with important feedbacks between the atmosphere and ocean in the subtropical Northern Hemisphere (section 3). And there is growing evidence that coupling between the atmosphere and oceans in the Northern Hemisphere extratropics contributes to the climate variability on decadal to multidecadal time scales.

This review paper contains a brief summary of the observational and modeling evidence for climate variability that is intrinsic to the coupling of the atmosphere and oceans in the tropics. The review is organized by phenomenon (e.g., the meridional modes, ENSO), and begins with the seasonal cycle in section 2 (because it is,

---

1 The interested reader can find more information on the pre-1960 observational studies that lead to an understanding of the Southern Oscillation in Wallace et al. (1998), Clarke (2008), and other readily available resources.

2 A glossary of acronyms is presented in the appendix.
in itself, a coupled atmosphere–ocean phenomenon and it is the background state that gives rise to ENSO) and the meridional modes in section 3. In section 4a we present the observations of ENSO, followed by a brief introduction to the essential uncoupled atmosphere and ocean dynamics that are relevant to ENSO in section 4b. We present the theory and further observations in support of ENSO as a true coupled atmosphere–ocean mode in sections 4c and 4d, respectively, and discuss the reasons for asymmetry between the warm and cold phases of ENSO in section 4e. A discussion of the sources of irregularity of ENSO is presented in section 4f, and a discussion of the response of ENSO to external forcing is presented in section 4g. Variability in the Indian Ocean is discussed in section 5.

We summarize in section 6a the observation and theoretical evidence for intraseasonal to interannual climate variability in the midlatitudes that is due to variability intrinsic to the midlatitude atmosphere and amplified by thermodynamic coupling to the ocean mixed layer. In section 6b we then present observations and modeling evidence for midlatitude variability on decadal and multidecadal time scales, in particular the Pacific decadal oscillation (PDO) and the Atlantic multidecadal oscillation (AMO)—both of which have been discovered within the past 30 years. We end with a summary of climate prediction based on these coupled modes of variability (section 7) and a brief discussion of problems that remain to be solved (section 8).

A note to the reader: in this review, we focus on phenomena that are fundamentally due to coupling between the atmosphere and ocean—by which we mean the existence, spatial patterns, and time spectrum of the phenomenon depend on processes in both the atmosphere and ocean. Mathematically, this idea of a coupled ocean–atmosphere phenomenon is described in Hirst (1986), who elegantly illustrated the presence of dynamical “modes” that only exist when the atmosphere and ocean are coupled. Here, we refer to “dynamic coupling” when ocean dynamics are fundamental for the coupled response, and “thermodynamic coupling” when the essence of the phenomenon can be understood without invoking ocean dynamics; ENSO is an example of the former and the meridional modes are an example of the latter. Coupled atmosphere–ocean phenomena are on seasonal and longer time scales. In some cases, coupled phenomena interact: for example, ENSO impacts the climate of the global tropics and the Western Hemisphere through teleconnections. For brevity, we do not focus on these teleconnections. We discuss the interaction between the coupled phenomenon only when the connections are essential (e.g., the impact of the Pacific meridional mode on ENSO). Phenomena that are intrinsic to the atmosphere but modified by their interaction with the ocean, or vice versa, are not covered in this review: review papers on the impact of the ocean on the atmospheric convection and on the Madden–Julian phenomenon—both phenomena intrinsic to the atmospheric—can be found in Hirons et al. (2018) and DeMott et al. (2016), respectively. Finally, we note that the field is rapidly expanding as observational and modeling capabilities expand and our theoretical understanding advances.

2. The seasonal cycle

In this section we introduce the climatology of the tropical oceans and atmosphere and briefly review the essential mechanisms that underlie the annual mean and seasonal cycle. In the Pacific Ocean and, to a lesser extent, the Atlantic Ocean, the essential physics appears to be due to coupling between the atmosphere and ocean.

a. The annual mean climatology

The seasonal cycle is mainly due to the variations in insolation associated with the obliquity of Earth’s axis of rotation with respect to its orbit, rendering two insolation maxima and two insolation minima on the equator each year. And so it is rather remarkable that the seasonal cycle of SST, wind, and upwelling along the equator in the tropical Pacific and Atlantic is predominately described by the annual harmonic rather than the semiannual harmonic. Indeed, the extrema in the seasonal cycle in the tropical Pacific and Atlantic are near the equinoxes (Figs. 8-1 and 8-2). The tropical cold season is in August–October (ASO) and features the strongest trade winds and upwelling in the deep tropics and a well-developed cold tongue with minimum SSTs; the warm season is in February–April (FMA) and features maximum SSTs, weakest trade winds and upwelling in the equatorial band, and relatively weak horizontal SST gradients (Fig. 8-1).

There is a strong east–west and hemispheric asymmetry in the annual averaged climatology in the tropical Pacific that is very similar to that in the tropical Atlantic. In the annual mean and throughout the seasonal cycle, the ITCZ is found north of the equator, a cold tongue exists in the eastern half of both basins, and the zonal average wind has a strong southeasterly component that crosses the equator in both the Atlantic and in the

3 Schlesinger and Ramankutty (1994) and Latif and Barnett (1996) were the first to report multidecadal variability in the North Atlantic Ocean and decadal variability in the North Pacific Ocean, respectively.
eastern half of the Pacific (Figs. 8-1 and 8-2). These perennial, asymmetric features are a fundamental ingredient for ENSO and for the coupled mode of interannual climate variability in the Atlantic that is ENSO-like (Zebiak 1993; see also section 4), which is referred to as the Atlantic Niño mode or Atlantic zonal mode, and so it is worth commenting on the sources of asymmetry in the annual mean climatology of these tropical basins.

In the zonal mean, the atmosphere transports heat southward across the equator, mainly by the zonally averaged Hadley circulation. Hence, the annual and zonal average intertropical convergence zone (ITCZ; defined as the centroid of precipitation; see Donohoe et al. 2013) is found ~1.7° north of the equator (Kang et al. 2008). However, inspection of Fig. 8-1 shows that the annual average location of the ITCZ in the Atlantic and eastern-central Pacific is much farther north—at about 7°N, and it only goes as far south as 5°N in the tropical warm season (FMA) in the Pacific (Fig. 8-3). The reason the ITCZ in the central and eastern Pacific remains in the Northern Hemisphere year-round is likely due to the presence of the Andes (Takahashi and Battisti 2007). The Andes block the midtropospheric westerly flow in the subtropics and cause dry air to descend in the subtropical southeast Pacific. This gives rise to an inversion, stratus clouds, and hence a year-round cooling of the subtropical southeast Pacific Ocean. In turn the cool air is advected by the trade winds into the central Pacific, creating a dry, cold wedge whose western boundary is the South Pacific convergence zone (Fig. 8-1).

With the ITCZ in the central-eastern Pacific confined to the Northern Hemisphere throughout the year, an annual averaged equatorial cold tongue and an east–west asymmetry in SST along the equator is ensured. Southerly winds blow across the equator all year round in the eastern half of the Pacific, causing upwelling (centered slightly off and south of the equator) and a tongue of cold water along the equator in the eastern Pacific (Mitchell and Wallace 1992). The latter drives an east–west sea level pressure (SLP) gradient and easterly mean winds along the equator (Lindzen and Nigam 1987; Battisti et al. 1999; Chiang et al. 2001; Back and Bretherton 2009a) that cause the thermocline to shoal in the eastern Pacific and deepen in the western Pacific. Bjerknes feedbacks (Fig. 8-4) and a thermocline that is shallow in the east and deeper in the western Pacific further amplify the east–west asymmetry in the ocean and strengthen the trades along the equator. As stated earlier, the annual mean asymmetry is essential for the ENSO mode to exist (see section 4).

The fundamental hemispheric asymmetry that keeps the annual average ITCZ to the north of the equator in the Atlantic (Fig. 8-1) is presumably the asymmetry in the geometry of Africa and South America (Privé and Plumb 2007b). The extraordinary surface heating of the Sahara in summer forces a monsoon circulation that is

---

**Fig. 8-1.** Climatology of the warm (FMA) and cold (ASO) seasons in the deep tropics. Shown are precipitation (contour interval 3 mm day$^{-1}$), SST (shading °C), and 10-m wind vector. Precipitation data are from the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) 1979–2017 (Xie and Arkin 1997a), SST data are from Hadley Centre Sea Ice and Sea Surface Temperature dataset (HadISST) 1870–2017 (Rayner et al. 2003), and wind data are from European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim) 1979–2017 (Dee et al. 2011).

**Fig. 8-2.** Seasonal cycle along the equator (2°S–2°N) in (top) SST (contour interval 1°C), (middle) zonal wind stress, and (bottom) meridional wind stress (contour interval 0.02 Pa, starting at ±0.02 Pa with negative contours dotted). Data sources are listed in the caption to Fig. 8-1.
barotropically and baroclinically unstable (Burpee 1972; see Wu et al. 2012, and references therein), spawning easterly waves across sub-Saharan Africa that (along with other synoptic disturbances) sum to make a well-defined Atlantic ITCZ in the Northern Hemisphere summer and early fall. The Southern Hemisphere summer monsoon is much weaker than the Northern Hemisphere monsoon, and so the annual mean ITCZ in the equatorial Atlantic is found to the north of the equator, with mean southeasterlies across the equator (Fig. 8-1).

b. The seasonal cycle of the Atlantic and Pacific ITCZs: The maritime monsoons

Poleward of ~6° latitude, the leading harmonic of insolation just off the equator is overwhelmingly annual. Thus, in the absence of other complicating factors (e.g., land), one would expect the ITCZs over the oceans to seasonally migrate from the Northern to the Southern Hemisphere and back again, lagging the local maximum in insolation at the solstices by about 3 months due to the thermal inertia of the ocean mixed layer, with the poleward limit of the ITCZs determined by symmetric instability (e.g., Emanuel 1995; Privé and Plumb 2007a). Were it not for the hemispheric asymmetry in the annual averaged ITCZ resides well north of the equator and (ii) there is a lesser seasonal cycle in the meridional position of the ITCZ (Fig. 8-3; Giese and Carton 1994; Chang and Philander 1994; Li and Philander 1996). The strength of the winter Hadley cell increases as the ITCZ moves farther into the Northern Hemisphere. Hence, the easterly winds south of the equator and the south-easterly wind that crosses the equator are strongest in ASO; the latter causes strong upwelling (just south of the equator) and gives rise to a strong SST minimum (a cold tongue) in the eastern tropical Pacific (Figs. 8-1 and 8-2; Mitchell and Wallace 1992). In turn, the zonal wind–SST gradient (Bjerknes) feedback increases the zonal trade winds along the equator and deepens and extends westward the cold tongue along the equator (Fig. 8-4; Chang and Philander 1994; Li and Philander 1996). By contrast, insolation north of the equator is weakest in

Fig. 8.3. Hovmöller diagram of the climatological SST and 10-m wind (1979–2017), averaged across the (left) Indian, (center) eastern half of the Pacific, and (right) Atlantic basin. SST is shaded (in °C) and precipitation is contoured (contour interval 2 mm day⁻¹). The wind vectors are relative to the maximum in each panel. Data sources are listed in the caption to Fig. 8.1.
November–January (NDJ) and so the water north of the equator is coldest in FMA. Hence, the ITCZ is found closest to the equator in FMA, and thus the meridional pressure gradient and cross-equatorial flow (and the upwelling it forces) is weakest in FMA. Similar processes act to create a dominant annual harmonic along the equator in the Atlantic.

The seasonal cycle in the near-equatorial tropical Indian Ocean is unlike that in the Pacific and Atlantic Oceans. Relative to the other two basins, winds in Northern Hemisphere summer monsoon oppose those of the winter monsoon, so that the annual average wind in the equatorial band is weak—including along the equator (Fig. 8-1). As a result, there is no cold tongue along the equator (Fig. 8-3). Indeed, the SST is relatively uniform north of 10°S throughout the Northern Hemisphere summer months. There is intense precipitation in the deep tropics (10°S–10°N) throughout the year. Heavy precipitation extends northward to 28°N in the Indian summer monsoon (Figs. 8-1 and 8-3) and heralds a different dynamics than is associated with ITCZs (see, e.g., Privé and Plumb 2007a; Bordoni and Schneider 2008; Boos and Kuang 2010).

3. Meridional modes

The emergence of our understanding of meridional mode variability has its seeds in early efforts to understand hydroclimate variations in the Nordeste region of Brazil and in the Sahel. Early investigations into hydroclimate variations in the Nordeste region focused on the influence of the Southern Oscillation, which at that time was not understood as a coupled component of ENSO. Indeed, as early as 1928, Sir Gilbert Walker noted that “the latest purpose to which [seasonal forecasting methods] have been directed is in connection with Ceará, a state in N.E. Brazil liable to terrible droughts” (Walker 1928). Later, severe drought in the Sahel from 1968 to 1974 prompted new investigations into the nature of hydroclimatic variations in that region [see Lamb (1978a,b), and references therein]. While these early studies identified coherent variations in large-scale conditions, they were constrained by the limited data availability, especially in the tropical and subtropical Atlantic (Landsberg 1975).

By the 1970s interest in predicting Nordeste and sub-Saharan precipitation variations, combined with the then available “seven decades of observations taken through the heroic and humble efforts of thousands of sailors and observers” (S. Hastenrath 2018, personal communication) led to a pioneering set of new studies that showed that the previously documented hydroclimatic variations were associated with coherent large-scale variations in both the ocean and atmosphere in the Pacific and Atlantic (Namias 1972; Hastenrath and Heller 1977; Markham and McLain 1977) [see also the study and extensive references in Moura and Shukla (1981)]. Hastenrath and Heller (1977) provided one of the first comprehensive analyses of large-scale oceanic and atmospheric conditions associated with Nordeste rainfall variations, showing that years of drought were associated with “an equatorward expansion of the South Atlantic, and a poleward retraction of the North Atlantic, high,” a concurrent weakening of the northeasterly and strengthening of the southeasterly trades, a northward shift in cloudiness and precipitation associated with the ITCZ, and positive (negative) SST anomalies north (south) of the equator. Moura and Shukla (1981) highlight the opposing SST anomalies on either side of the equator, and Hastenrath and Greischar (1993) established a physical causal mechanism between those SST anomalies and the large-scale conditions, featuring the hydrostatic adjustment of the lower troposphere to the SST anomalies (Lindzen and Nigam...
At the same time, similar large-scale SST variations were implicated in variations in Sahel precipitation (Fig. 8-5; Folland et al. 1986; Lough and Lough 1986; Hastenrath 1990). The causal mechanism outlined in Hastenrath and Greischar (1993) was generally confirmed in a variety of subsequent observational and modeling studies (Ruiz-Barradas et al. 2000; Chang et al. 2000; Chiang et al. 2001; Chiang 2002). Still lacking, however, was an understanding of the cause of the SST anomalies in the first place, which was emphasized in the closing of Hastenrath and Greischar (1993): “While the novel aerological data source has thus served to identify pivotal atmospheric processes in the climate dynamics of the tropical Atlantic, the mechanisms operative in the origin and maintenance of prominent hydrospheric anomalies are less well understood.”

Improved data availability in the tropical Atlantic led to the understanding that these coordinated SST and wind anomalies were due to a coupled atmosphere–ocean mode in the tropical Atlantic. Observational studies that focused on SST alone tended to reproduce the “dipole” structure (Servain 1991) formed by the northern and southern tropical Atlantic SST anomalies on multidecadal time scales, but emphasized the lack of interhemispheric coherence between the two centers of action on interannual to decadal time scales (Houghton and Tourre 1992; Mehta and Delworth 1995; Xie and Tanimoto 1998; Enfield and Mestas-Núñez 1999). In contrast, analyses that combined ocean and atmosphere variables—especially low-level meridional winds—tended to retain the dipolar structure on all time scales (Nobre and Shukla 1996; Chang et al. 1997; Ruiz-Barradas et al. 2000; Chiang and Vimont 2004). These results are generally interpreted to imply that the key element of the coupled variability is the cross-equatorial SST gradient, and that a dipolar structure tends to maximize that gradient (Chiang and Vimont 2004). Recent theoretical work, however, calls that interpretation into question (Martinez-Villalobos and Vimont 2017).

By the late 1990s, two (not necessarily distinct) explanations for the tropical Atlantic variability emerged. The first described the variability as a response to stochastic variability in the subtropical trade winds—some of which is related to the North Atlantic Oscillation—or from ENSO-related atmospheric teleconnections (Curtis and Hastenrath 1995; Nobre and Shukla 1996; Xie and Tanimoto 1998; Giannini et al. 2000; Czaja et al. 2002). Stochastic variations in subtropical trade winds impart an SST anomaly on one side of the equator that causes a tropical atmospheric response through low-level adjustment to the cross-equatorial SST gradient. Alternatively, variations in cloudiness, or dust advected over the Atlantic, can force meridional mode variations through changes in shortwave radiative fluxes (Tanimoto and Xie 2002; Evan et al. 2009, 2011). The second explanation ascribed the tropical Atlantic variability to be a coupled atmosphere–ocean mode that features a self-sustaining oscillation by way of a thermodynamic feedback between the surface wind, wind-induced evaporation, and SST (the wind–evaporation–SST (WES) feedback;
Xie and Philander 1994; Chang et al. 1997]. Model analysis confirmed the existence of a self-sustaining mode of variability in simple coupled ocean–atmosphere models that include a WES feedback (Xie 1997b; Chang et al. 1997; Zhou and Carton 1998; Xie 1999; Kossin and Vimont 2007). Of course, these studies recognized that the “forced” versus “internal” interpretations were not mutually exclusive. A variety of observational and modeling studies demonstrated that both external forcing and a discernible WES feedback operate in the evolution of tropical Atlantic variability (see, e.g., Xie and Tanimoto 1998; Okumura et al. 2001; Czaja et al. 2002; Kushnir et al. 2002b; Frankignoul and Kestenare 2005, among others). These studies showed that subtropical trade wind variations, often associated with remote forcing from ENSO or the North Atlantic Oscillation, generate SST anomalies in the subtropical North Atlantic through wind-induced latent heat fluxes (forcing); the resulting SST anomalies then induce their own circulation including a positive WES feedback in the near-equatorial region (the positive feedback is limited to the northern deep tropics; there no evidence of positive WES feedback south of the equator; Amaya et al. 2017).

Despite growing evidence of a coupled mode of variability intrinsic to the tropical ocean–atmosphere system, there was still considerable debate as to the nature of the dipole-like mode of variability in the Atlantic, and a general acceptance that the variability in question only existed in the Atlantic. In 2002 John Chiang approached Dan Vimont and noted that the chain of events associated with cross-equatorial variability in the tropical Atlantic outlined in Czaja et al. (2002) and Kushnir et al. (2002b) bore a strong resemblance to the “seasonal footprinting mechanism” (Vimont et al. 2001, 2003a,b) in the Pacific. Chiang and Vimont (2004) applied maximum covariance analysis (MCA) to observed SST and low-level winds to show that indeed the tropical Pacific contains a structure of covariability that is strongly analogous to that in the Atlantic (see Fig. 8-5). Following the lead of Servain et al. (1999), they referred to the analogous modes of variability as “meridional modes.”

Chiang and Vimont (2004) showed that in both the Pacific and Atlantic, the meridional modes feature an anomalous meridional SST gradient that maximizes near the position of the climatological ITCZ, a meridional shift of the ITCZ toward the warm SST anomalies, and that both the Pacific and Atlantic meridional modes (PMM and AMM, respectively) are forced by analogous intrinsic modes of midlatitude atmospheric variability in the Northern Hemisphere of their respective basins (the North Pacific Oscillation (NPO; Linkin and Nigam 2008) and North Atlantic Oscillation (NAO)] and appear to involve subsequent evolution equatorward of the original forcing through a WES feedback. Further, they used atmospheric model simulations forced by meridional mode SST structures to show that the subtropical component of the covarying wind anomalies was independent of SST, while the tropical component of the wind was indeed a response. The strong similarity between the meridional mode structure in the Pacific and Atlantic showed that meridional modes are ubiquitous features of the coupled tropical ocean–atmosphere system. Since then, meridional mode–like variability has been identified in the Indian Ocean (Wu et al. 2008) and in the South Pacific (Van Loon and Shea 1985; Zhang et al. 2014).

a. Dynamics

A critical component of tropical meridional mode variability is the WES feedback acting alongside changes in the cross-equatorial SST gradient (CESG). In an excellent review by Xie and Carton (2004), they summarize the tropical Atlantic variability as follows: “A positive anomalous CESG sets up an anomalous southward pressure gradient in the atmospheric boundary layer … inducing southerly cross-equatorial winds that decelerate the easterly trades north of the equator because of the Coriolis effect… [These] weakened trades north of the equator reduce surface evaporation, thereby acting to strengthen the initial CESG. South of the equator, the south-easterly trades accelerate, increasing surface evaporative cooling and the northward CESG.” The existence of the WES feedback has been investigated in a variety of model and observational studies (Chang et al. 2000; Sutton et al. 2000; Okumura et al. 2001; Czaja et al. 2002; Chiang and Vimont 2004; Frankignoul and Kestenare 2005; Smirnov and Vimont 2011; Amaya et al. 2017), and the studies generally support the existence of a positive feedback in the deep tropics (within about 15° of the equator; Chang et al. 2000; Lin et al. 2008), but not in the subtropics. Additional processes have been shown to impact tropical Atlantic meridional mode variability, including Ekman currents (Xie 1999; Kushnir et al. 2002b), mean ocean currents (Chang et al. 1997, 2001; Seager et al. 2001; Kushnir et al. 2002b), and short-wave cloud feedbacks (Okumura et al. 2001; Tanimoto and Xie 2002; Evan et al. 2013).

The essential physics of meridional mode growth and propagation is captured by simple coupled ocean–atmosphere models of the tropical ocean–atmosphere system (Xie 1999; Wang and Chang 2008a; Vimont 2010; Martinez-Villalobos and Vimont 2017). Vimont (2010) coupled a steady Gill–Matsumo–type atmosphere model (Gill 1980; Matsumo 1966) to a motionless “slab” ocean model to investigate the physics of meridional model propagation and growth. The key coupling process included in these simple models is the WES process, in which
variations in zonal wind generate anomalies in evaporation, and hence affect SST: in a region of easterly mean trade winds, a westerly wind anomaly results in a relaxation of the climatological wind speed, a reduction in evaporation, and hence a warming to the ocean. While meridional mode–like structures can be obtained as free eigenmodes (eigenvectors) of the linearized dynamical system (Xie 1999), coupling makes the dynamical matrix nonhermitian. As such meridional modes can experience transient growth via nonnormal interactions of the eigenmodes (Penland and Matrosova 1998; Chang et al. 2004a,b; Wang and Chang 2008b; Vimont 2010; Martinez-Villalobos and Vimont 2017). In such a system, the pattern that experiences the greatest transient growth strongly resembles the observed meridional mode (Fig. 8-6), with a dipolar SST structure that is antisymmetric about the equator, a surface pressure gradient force toward the warmer hemisphere, and equatorial surface winds that blow toward the anomalously warm hemisphere (in the antisymmetric case).

A key finding in the analysis of these simple models is the growth and propagation mechanism of tropical meridional modes (Vimont 2010; Martinez-Villalobos and Vimont 2017). Meridional modes experience equatorward and westward propagation (Liu and Xie 1994; Vimont et al. 2009; Amaya et al. 2017), which is a consequence of the phasing of the zonal wind anomalies with respect to the SST anomalies. Taking the positive SST anomaly north of the equator in Fig. 8-6b as an illustrative example, a given warm surface temperature anomaly will generate a cyclonic low-level circulation that includes westerly wind anomalies (a relaxation of the climatological trades) to the south of the original SST anomaly. The reduced wind reduces evaporation and causes the warm SST anomaly to propagate equatorward. Martinez-Villalobos and Vimont (2017) show that the westward propagation is consistent with the westward group velocity of the (long) coupled atmospheric equatorial Rossby waves that are coupled with the evolving meridional mode SST anomalies.

Growth of the meridional mode structure results from the spatial covariance between the wind-induced evaporation and the anomalous SST (Xie 1999; Vimont 2010). While this spatial covariance is often described as a consequence of the reversal of the Coriolis parameter across the equator, Vimont (2010) showed that in the case of the meridional mode it is related to the meridional offset between the atmospheric response to heating and the SST (which is responsible for the heating) itself. This meridional offset is due to the variation of the Coriolis parameter with latitude, and hence reversal of the Coriolis parameter is not a necessary condition. Heating generates stretching in the horizontal vorticity equation through heating-induced convergence.

But this stretching only acts as a source of vorticity when it occurs in the presence of background vorticity (planetary vorticity in this case). As such, the meridional variation of the Coriolis parameter with latitude implies that this stretching is most effective near the poleward flank of the SST anomaly; hence, the atmospheric response is shifted poleward (Fig. 8-6). This poleward offset means that the domain-integrated WES feedback—which requires a collocation of winds and SST—is positive. This also means that the WES feedback is most effective where the background vorticity is changing the most rapidly: near the equator.

![Fig. 8-6. Growth and propagation of a meridional mode-like structure using a Gill–Matsuno model coupled to a slab ocean model, as in Vimont (2010). (a) An idealized initial SST structure at day 0, (b) day 60 of the integration, and (c) day 120 of the integration. Shading depicts SST, vectors low-level wind, and contours low-level geopotential. Solid contours denote positive geopotential anomalies, and dashed contours denote negative geopotential anomalies. Units are arbitrary, but consistent throughout the panels. Note the equatorward and westward propagation of SST anomalies, and the increased amplitude (growth) of SST anomalies throughout the integration.](https://example.com/figure86.png)
b. Impacts

In addition to hydroclimate variations in the Nordeste and Sahel, meridional modes impact both weather and climate conditions in the Pacific and Atlantic. In the Pacific, meridional modes play a key role in the seasonal footprinting mechanism [see section 4f(1)], and as such provide a link between mid- to high-latitude atmospheric variations and tropical Pacific ENSO (Vimont et al. 2001, 2003a,b; Chang et al. 2007; Alexander et al. 2010; Larson and Kirtman 2013). This is especially true for central Pacific ENSO events (Vimont et al. 2014; Lin et al. 2015). Similar influences have been suggested from the Southern Hemisphere (Zhang et al. 2014; You and Furtado 2017), though the interpretation of the South Pacific meridional mode as a potential influence on ENSO is still in question (Larson et al. 2018). Vimont (2005) shows that the Pacific meridional mode contributes to the spatial structure of decadal variability in the Pacific (Zhang et al. 1997) through its influence on ENSO, which was later shown through modeling studies (Di Lorenzo et al. 2015).

An important consequence of AMM variability is its relationship with tropical cyclone activity. Vimont and Kossin (2007) and Kossin and Vimont (2007) show that the AMM plays an important role in tropical cyclone activity through its impact on a number of environmental conditions that all cooperate in their influence on tropical cyclone activity (Smirnov and Vimont 2011). Patricola et al. (2014) show that the AMM and ENSO work together to influence Atlantic tropical cyclone activity. The PMM has also been shown to influence Pacific tropical cyclone activity (W. Zhang et al. 2016; Zhang et al. 2017; Zhan et al. 2017).

4. ENSO

According to the Web of Science, more than 4800 scientific papers have been published in the past 20 years with either ENSO or El Niño in the title (for reference, over the same time period “synoptic” appears in less than 1900 papers). Here we limit our discussion to focus on the fundamental aspects of ENSO and highlight the advances in our understanding of ENSO over the past 20 years—advances that complement and refine the earlier (and still relevant) papers that summarize our understanding of ENSO from models and observations. In addition to the cited literature, several books on ENSO provide useful material, including those by Philander (1990), Clarke (2008), and Sarachik and Cane (2010).

a. Anatomy of ENSO from observations

The essential signature of ENSO as seen in SST, surface wind, SLP, and precipitation is summarized in the reviews by Rasmusson and Carpenter (1982), Wallace et al. (1998), Larkin and Harrison (2002), and Chen and Wallace (2015) and is shown in the top three panels of Fig. 8-7. Warm (El Niño) events are characterized by greater-than-normal SST in the equatorial eastern and central tropical Pacific, relaxed trade winds on the equator in the west-central Pacific, and an eastward shift in the climatological mean precipitation from the Maritime Continent to the central Pacific. Note that the maximum zonal wind (stress) anomalies are collocated with the maximum zonal gradient in anomalous SST and with the maximum precipitation anomaly—although the latter is due to anomalous moisture flux convergence primarily associated with the meridional wind anomalies [see section 4b(1)]. The spatial structure of the cold phase of ENSO is qualitatively similar to that in the warm phase, although there are quantitative differences in the location of the SST and precipitation (see, e.g., Hoerling et al. 1997).

Although gross aspects of the evolution of warm (El Niño) events are common to almost all warm events, the spatiotemporal evolution of each warm event differs in detail [see, e.g., the discussion in Rasmusson and Carpenter (1982) and Wallace et al. (1998) and numerous studies in the past two decades]; the same is true for cold (La Niña) events. For example, some warm events start in the far eastern basin and quickly propagate westward,
while others start in the central basin and quickly propagate eastward; still other events show little propagation. Similarly, although warm events tend to last 12 months, the duration of a warm event varies from 9 months to almost 2 years. These differences are often referred to in the literature as different “flavors” of ENSO or ENSO diversity (Trenberth and Stepaniak 2001; Capotondi et al. 2015b); the essential physics that grow and decay all warm and cold events are similar, but minor differences in the mixture of processes (e.g., the mix of process affecting SST; Zebiak 1984; Hirst 1986; Battisti 1988; Boucharel et al. 2015) give rise to different flavors of ENSO.

There are many different indices of the state of ENSO. The most common indices are SST anomalies averaged over selected regions in the equatorial Pacific. The top panel in Fig. 8-8 shows the time series of Niño-3.4 (the SST anomalies averaged from 5°S to 5°N and from 170°E to 120°W), which is a good index of the large-scale state of the tropical Pacific Ocean. Significant warm events are found in 1965/66, 1972/73, 1982/83, 1987/88, 1997/98, and 2015/16; lesser events are found in other years. Warm events typically last 9–18 months and are followed by weaker cold events that typically last longer than warm events, particularly in the eastern half of the Pacific. Averaged over seasonal and longer time scales, Fig. 8-8 shows that Niño-3.4 is highly correlated with the Southern Oscillation index (SOI) at zero lag ($r = 0.93$)—an indication that ENSO is born of coupling between the atmosphere and ocean. Hence, Niño-3.4 and several other common indices [e.g., SOI, trans-Niño index (TNI), oceanic Niño index (ONI), and cold tongue index (CTI); see Trenberth and Stepaniak 2001] are also good indices of the state of the atmosphere in the tropical Pacific (see, e.g., the time series of precipitation in the central Pacific in Fig. 8-8).

Dynamical changes in the depth of the thermocline in the tropical Pacific are fundamental for ENSO (see section 4c). Figure 8-9 (from Chen and Wallace 2015) shows the leading empirical orthogonal function (EOF) of the three-dimensional ocean temperature in the tropical Pacific. The time series is correlated with Niño-3.4 at 0.9. The climatological thermocline, generally taken to be the depth of the 20°C isotherm in the equatorial tropics, is deeper in the eastern Pacific and shallower in the eastern Pacific (due to the annual mean easterlies on the equator). The top panel in Fig. 8-9 thus shows that during El Niño, the thermocline is anomalously deep in the eastern Pacific and anomalously shallow in the western Pacific. In the eastern Pacific (bottom right), temperature anomalies are centered on the equator and confined to a few degrees about the equator—a signature of the oceanic Kelvin mode response to westerly wind anomalies along the equator in the west-central Pacific. In the western Pacific (bottom left), temperature anomalies extend farther poleward and the structure is very similar to what is expected from the gravest equatorially trapped Rossby modes. The Kelvin and Rossby signatures are also clearly seen in Fig. 8-7, which shows the steric sea level height (including the correlated salinity changes) associated with the warm phase of ENSO. Altogether, the three-dimensional
oceanic temperature structure that is correlated with ENSO is consistent with an adiabatic dynamical response to changes in the zonal wind stress forcing on the equator in the west-central Pacific (shown in the top two panels of Fig. 8-7).

Finally, there is a telltale systematic temporal evolution in the thermocline during ENSO cycles, illustrated by a Hovmöller diagram (Fig. 8-10). Well-developed warm phases of ENSO (El Niño) are seen in 1982/83, 1987/88, and 1997/98—each of which is followed by a cold phase (a La Niña). The SST anomalies are confined to the east of the date line and evolve roughly synchronously. And although the wind stress anomalies are in phase with the SST anomalies, the thermocline is not in equilibrium with the wind stress. There is a clear slow eastward propagation of thermocline anomalies in the equatorial band; this is the signature of ocean memory and is the reason that warm events are followed by cold events in a cycle that lasts 3 or 4 years (ranging up to 7 years), and it is the hallmark of the ENSO mode (see section 4c).

There is a long literature on the far-field impacts of ENSO on climate. The most notable impacts are on precipitation and temperature throughout the global tropics (Klein et al. 1999), and significant extratropical temperature and precipitation anomalies that are communicated by atmospheric planetary waves that extend to western Antarctica and South America, and in wintertime across the North Pacific and North and Central America. See Trenberth et al. (1998) and Yeh et al. (2018) for reviews of the far-field impacts of ENSO and the dynamics of these teleconnections.

ENSO also has an impact on the extratropical ocean. Atmospheric teleconnections to the North and South Pacific Ocean leave an imprint on ocean SST through their impact on the local turbulent energy fluxes; this physics is referred to as “the atmospheric bridge” (Alexander 1992; Lau and Nath 1996; Alexander et al. 2002). As anticipated by theory (Moore 1968) and seen in tide gauges (Enfield and Allen 1980; Clarke and Liu 1994), coastal oceanic Kelvin waves carry the signature of ENSO from the eastern tropical Pacific boundary to the midlatitude oceans along the west coast of the Americas as well as into the Indian Ocean (see Clarke and Liu 1994)—evident in the strong correlation between ENSO and the sea level difference between Perth and San Francisco shown in Fig. 8-8.

b. Uncoupled atmosphere and ocean dynamics essential for seasonal to interannual coupled variability in the tropics

In a pioneering paper, Matsuno (1966) provided wave form solutions to the time-dependent, linearized, shallow water equations on an equatorial \( \beta \)-plane that are relevant to both the atmosphere and ocean. These solutions feature an eastward semigeostrophic Kelvin wave, westward Rossby waves, a mixed Rossby gravity wave, and eastward and westward propagating interio-gravity waves. Many studies have demonstrated that free (unforced) transient features with these structures are ubiquitous in the troposphere, stratosphere, and upper ocean. For example, Maruyama (1968) and Kousky and Wallace (1971) document stratospheric mixed Rossby gravity waves and Kelvin waves, respectively, that were later
shown to be important for driving the stratospheric quasi-biennial oscillation (Holton and Lindzen 1972). In the troposphere, mixed Rossby gravity and Kelvin waves organize convection, along with other modes that are driven by convection and have no dry analogs in Matsuno’s system, such as easterly waves and the Madden–Julian oscillation (see Kiladis et al. 2009, and references therein).

To understand ENSO requires an understanding of the response of the tropical atmosphere to the distribution of SST, and of the dynamic and thermodynamic response of the ocean to surface wind stress forcing: these are briefly outlined in this section, followed by a summary of the coupled ENSO mode. These tools help us to understand the evolution of ENSO shown in section 4a and form the foundation for the intermediate coupled atmosphere–ocean models upon which the theory of the ENSO mode (see section 4c) is based. As we shall see, oceanic Kelvin and Rossby waves are issued in response to zonal wind stress anomalies and comprise the “ocean memory” that is fundamental to ENSO and the ENSO mode (see section 4c).

1) THE RESPONSE OF THE ATMOSPHERE TO DIABATIC HEATING

Atmosphere and ocean indices of ENSO show that the ENSO mode evolves slowly—on time scales of more than 2–3 months. On these time scales, the atmosphere is nearly in a statistical equilibrium and forced by the distribution of diabatic heating (the adjustment to the equilibrium state is by transients, however), and for this reason Gill (1980) explored the steady solutions to Matsuno’s equations, assuming frictional damping and forcing by a prescribed, stylistic spatial pattern of heating with a simple, vertical structure in the troposphere (sinusoidal in pressure, with maximum in midtroposphere). His solutions shared much in common with the observed tropical circulations in response to condensational heating.
Before the Gill model could be coupled to an ocean, an additional problem had to be solved: how to determine the distribution of heating. If heating is due to condensation in deep convection and, in turn, the condensation is supplied by moisture convergence, then in equilibrium the winds generated by the heating must be consistent with the convergence of water required to balance the condensation. Still retaining a single vertical mode, Zebiak (1986) solved this problem with an iterative solution to Gill’s model forced by turbulent heat fluxes at the atmosphere–ocean interface, and he obtained somewhat realistic surface wind anomalies when the Gill model was forced by canonical SST anomalies associated with the warm and cold phases of ENSO.

Lindzen and Nigam (1987) took an entirely different approach. They assumed that the winds in the boundary layer were entirely due to pressure gradients that develop in response to boundary layer temperature gradients. In their model, boundary layer convergence would provide the fuel for deep convection aloft, but they assumed that the latter did not further affect the boundary layer flow. Neelin (1989) showed that the Lindzen and Nigam (1987) model was mathematically equivalent to Zebiak’s form of the Gill model, and so it was not surprising that the two models gave similar solutions—albeit with entirely different explanations (see also Wang and Li 1993; Battisti et al. 1999).

Chiang et al. (2001) resolved the apparent incongruity of the Lindzen–Nigam and Zebiak–Gill models by fitting the observed three-dimensional wind anomalies in the tropical Pacific to a linearized primitive equation model of the atmosphere that included an implicit boundary layer. They showed that the boundary layer meridional wind and its convergence are primarily driven by the low-level pressure gradient associated with SST gradients. In turn, the boundary layer convergence fuels an upper-layer condensational heating that greatly affects the zonal winds in the boundary layer, but it has a modest contribution to boundary layer convergence compared to the meridional winds driven by the SST gradients (see Fig. 8-11). Back and Bretherton (2009a,b) came to the same conclusions as Chiang et al. (2001) using a mixed-layer boundary layer model and utilizing reanalyses data. Adames and Wallace (2017) also confirmed Chiang’s findings using a longer dataset and entirely empirical methods (see Fig. 8-11). Since most intermediate coupled atmosphere–ocean models (see next section) use some variant of the Zebiak formulation, it is somewhat fortuitous that solutions to this form of the Gill model conform to the physics operating in nature. More specifically, it is
fortunate that these models give realistic zonal wind stress along the equator, because the latter is what is important for the two essential processes acting to form the ENSO mode: the Bjerknes feedback and the adiabatic ocean adjustment to wind stress changes (see section 4c).

2) THE RESPONSE OF THE OCEAN TO WIND STRESS FORCING

The ENSO mode and the theory of ENSO are based on models and eigensolutions of the coupled atmosphere–ocean system in the tropical Pacific. These intermediate coupled models typically employ equations describing the adiabatic adjustment of an ocean basin to wind stress forcing and assume one or two vertical baroclinic modes (see section 4c). Numerous studies have forced ocean general circulation models and linearized, adiabatic ocean models of the tropical Pacific with the observed history of wind stress. These studies find a reasonably good agreement between the observed and simulated thermocline and sea level variations, and between observed and calculated thermocline and sea level variations using the linearized, adiabatic equatorial wave theory (see, e.g., Busalacchi et al. 1983; Richardson and Philander 1987; Bigg and Blundell 1989; Kessler 1990). The observed thermocline and sea level variations are reasonably well approximated by the sum of long, non-dispersive westward propagating Rossby waves and an eastward propagating Kelvin wave, with appropriate eastern and western boundary conditions (Moore 1968; Cane and Sarachik 1977; Clarke 1983, 1991; Du Penhoat and Cane 1991; Clarke 1992).

Of course, the atmosphere is forced by surface energy fluxes that are determined by the distribution of SST, which in turn is also affected by upwelling, mixing, and horizontal advection—processes that are affected by thermocline variations in various degrees. Many studies have shown that intermediate ocean models and general circulation models, when forced by the observed history of wind stress, reproduce reasonably well the observed interannual variability in SST (e.g., Seager 1989; Stockdale et al. 1998b).

Neelin and Jin (1993) showed that in the longwave limit of nondispersive Rossby waves, the free eigen-modes of a linearized bounded tropical ocean basin of the size of the Pacific are inconsistent with the thermocline evolution in an ENSO cycle. For the same geometry, however, forcing by a periodic interannual zonal wind stress features a thermocline evolution that is surprising similar to that observed (e.g., Cane and Sarachik 1981; Neelin and Jin 1993). This result suggests that ENSO can only come about through a coupling between the atmosphere and ocean in the tropics, as is implied from the strong correlations between atmosphere and ocean indices of ENSO, as seen in Fig. 8-8.

3. Modeling and theory of ENSO

A central achievement of the TOGA program was the development of a theory for ENSO that has stood the test of time. Here we present a summary of the essential physics of the ENSO mode that is pertinent to nature and refer the interested reader to the post-TOGA review paper by Neelin et al. (1998) for a more complete historical background. Research in the past 20 years or so has focused on sources of irregularity in ENSO, evaluation of ENSO in the climate models used in the Intergovernmental Panel on Climate Change (IPCC) assessment reports, and the seasonal-to-interannual predictability of ENSO (see sections 4f, 4g, and 7).

1) THE ENSO MODE

The large-scale El Niño–Southern Oscillation phenomenon is unique in that it is intrinsically due to coupling between the atmosphere and ocean in the tropical Pacific: uncoupled atmosphere models forced by the seasonal cycle of SST do not produce realistic Southern Oscillation, and uncoupled ocean models forced by the seasonal cycle of surface wind stress do not produce pan-Pacific El Niño and La Niña events. Among the first models that simulated realistic ENSO cycles were “intermediate coupled models” of the upper ocean in the tropical Pacific coupled to an atmosphere model that was in equilibrium (or nearly so) with the SST produced by the ocean model (see section 4b), including the coupled models developed by Zebiak and Cane (1987) and Schopf and Suarez (1988). The ocean component of the intermediate coupled models incorporates the myriad essential processes that affect SST (i.e., parameterizations for the impact of upwelling and vertical mixing, surface fluxes, advection) and linear or nonlinear adiabatic dynamics of the upper few hundred meters of ocean; these models contained the same physics as the models used to successfully hindcast the SST and sea level variations in ENSO [see section 4b(2)]. The atmosphere component of the intermediate models was usually some form of the Gill model, with atmospheric forcing parameterized in terms of the SST anomalies [see section 4b(1)].

Zebiak and Cane (1987) constructed an intermediate coupled model that simulated anomalies about prescribed climatological surface winds, SST, and ocean currents. The Zebiak–Cane model was the first model to simulate realistic ENSO variability—yearlong warm events followed by yearlong cold events and with a

---

5 Short Rossby waves have slow eastward group velocity and thus do not propagate very far from the source of forcing or away from the western boundary.
spatiotemporal evolution in SST, zonal wind stress, and thermocline displacement that were reasonably similar to that observed. A year later, Schopf and Suarez (1988) developed a slightly different intermediate coupled model that also displayed ENSO cycles that were reasonably realistic (Fig. 8-12). Schopf and Suarez’s analysis, as well as Battisti and Hirst’s analysis of a variant of the Zebiak-Cane model (Battisti 1988; Battisti and Hirst 1989), showed that the essential physics of ENSO is contained in two competing processes: (i) atmosphere–ocean coupling—the Bjerknes feedback—that relates the strength of the zonal wind anomalies in the equatorial band of the central Pacific to the amplitude of the SST anomalies localized to the eastern half of the Pacific and (ii) the slow dynamical adjustment of the equatorial ocean by Kelvin and Rossby waves forced by changes in zonal wind stress—the so-called ocean memory. The Bjerknes feedback is “localized” in the central-eastern Pacific; there is little change in SST in the western Pacific because horizontal gradients in SST are weak, and the thermocline is deep.

The interannual time scale for the ENSO mode results because these two processes—which have time scales of about 6 and 9 months, respectively—oppose each other, and because the impact of the ocean dynamical adjustment on SST is confined to the eastern half of the Pacific and not felt for 6–9 months after the initial changes in wind stress (i.e., it takes 6–9 months for Rossby waves issued by zonal wind stress changes in the central Pacific to reach the western boundary and return to the central

FIG. 8-12. Dynamical evolution of ENSO in the equatorial ocean in an intermediate coupled atmosphere–ocean model of the tropical Pacific, traced by the gray and white swaths from the lower left to the upper right. Kelvin (\(\eta_K\)) and Rossby (\(\eta_R\)) signals are highlighted by way of averaging sea level along latitude bands in which sea level variations are predominantly due to the gravest (fastest) Rossby mode (between 5° and 7° north and south) and Kelvin mode (the equator). Starting from an arbitrary phase, an easterly zonal wind anomaly \(U\) associated with the cold phase of ENSO in the center panel in year 0–2.5 generates a positive Rossby wave packet \(\eta_R\) in the far left panel that is seen to propagate slowly westward from the central Pacific to the western boundary, where it issues a positive (downwelling) Kelvin wave packet \(\eta_K\) in the center-left panel that propagates eastward, causing an El Niño condition in year 4–5 and reversing the wind anomaly in the central Pacific (center panel). In turn, the easterly wind anomaly generates a negative Rossby wave in the central Pacific that propagates westward (center-right panel), and issues in year 6 a negative (upwelling) Kelvin wave packet in year that reaches the central Pacific in year 6–7 that causes a La Niña and easterly wind anomalies, completing the ENSO cycle. [From Schopf and Suarez (1988).]
and eastern Pacific in the form of a Kelvin wave, where it can further affect SST. The physics of ENSO cycles described by Schopf and Suarez (1988) and Battisti and Hirst (1989) were coined “the delayed oscillator” theory for ENSO (Suarez and Schopf 1988).\footnote{Schopf and Suarez (1988) and Battisti and Hirst (1989) developed toy models to illustrate the essential physics of the ENSO cycle that were later given the moniker “delayed oscillator” model for ENSO. Jin (1997) also developed a toy model for the ENSO mode that he called “the recharge oscillator.” These two toy models express the same physics (the delayed oscillator casts the ocean memory in terms of wave dynamics, while the recharge oscillator casts the ocean memory in terms of Sverdrup balance). Given the parameter values used in the recharge oscillator, however, these two toy models are one in the same. Both toy models are very crude representations of the ENSO mode and have little value beyond a heuristic representation of essential ENSO physics.}

The essential dynamics of ENSO appears to be governed by linear dynamics and thermodynamics. Battisti and Hirst (1989) linearized the coupled Zebiak–Cane model about the climatological and annual averaged basic state of the model; schematically \( \dot{x} = Mx \), where \( x \) is the vector containing the state variables and \( M \) is the linearized physics matrix describing the dynamics and thermodynamics of the atmosphere and upper equatorial Pacific ocean. They showed that the structure of the warm and cold ENSO events and the period of ENSO in the Zebiak–Cane model are well described by the first (unstable) eigenmode of \( M \) (Fig. 8–13). In particular, the temporal and spatial evolution of SST, wind stress, and thermocline in the ENSO eigenmode are similar to those in observations [cf. Figs. 8–10 and 8–13 (top); Figs. 8–7c and 8–13 (bottom left)], albeit with the center of action shifted too far east (this bias is remedied when more realistic mean states are used in the formulation of \( M \); see Atwood et al. 2017).

The anatomy of the ENSO mode in the Zebiak–Cane model is very similar to that simulated by many other intermediate and full climate models. Nonlinearity acts mainly to bound the growing eigenmode and create a finite amplitude ENSO cycle; it does not affect the spatiotemporal structure of the warm and cold events (Fig. 8–13), or qualitatively effect the period of the ENSO cycle (see, e.g., Battisti and Hirst 1989; Jin et al. 1996). As in observations, tropical ocean dynamics are essential to the ENSO eigenmode, and the spatial and temporal evolution of SST anomalies are due to a rich mixture of ocean dynamics and surface fluxes (e.g., Zebiak 1984; Hirst 1986; Battisti 1988; Boucharel et al. 2015).

2) FURTHER THEORETICAL RESULTS RELEVANT TO THE ENSO MODE

That ENSO is a true eigenmode of the coupled atmosphere–ocean system in the tropical Pacific is also supported by eigenanalyses of the empirically derived dynamical matrix \( M \), estimated from a fit of the observed history of SST anomalies in the tropical Pacific to a linear inverse model (LIM). In a pioneering study, Penland and Sardeshmukh (1995) showed that the leading eigenmode of the empirically derived \( M \) has the SST footprint of ENSO with a period of 3.8 years. Similar results are found in other studies (see Newman et al. 2011a, and references therein). Unlike the ENSO mode in the Zebiak–Cane model, the ENSO mode determined from the empirically derived \( M \) is asymptotically stable—as it must be when the dynamical operator is estimated from a geophysical time series. Thompson and Battisti (2000) performed an eigenanalysis of the operator \( M \) from the Zebiak–Cane model but with more realistic parameters and retaining the seasonal cycle in the basic state (the basic state was updated with observations that were not available in the 1980s when the original Zebiak–Cane model was constructed). With the updated parameters and climatological seasonal cycle basic state, the leading eigenmode features very similar spatial structures to those observed in ENSO, and to those obtained from the leading eigenmode of the original model. The ENSO eigenmode has an average period of 3.9 years, and it is asymptotically stable.

As was evident from the earliest studies of the coupled atmosphere–ocean system in the tropics (Hirst 1986, 1988), gross aspects of the ENSO are predicated on there being east–west asymmetries in the annual averaged climatological SST (i.e., a cold tongue in the eastern basin) and thermocline, as well as in the zonal and meridional wind stress (see Fig. 8–1 and the discussion in section 2). Sensitivity studies show that changes in the annual averaged climatology greatly affect the stability and structure of the ENSO mode (e.g., Zebiak and Cane 1987; Battisti 1988; Battisti and Hirst 1989; Neelin et al. 1998, and references therein). Furthermore, for the same mean state climatology, changes in the strength in any of the myriad various process that affect SST affect the detailed evolution of warm or cold events—for example, whether warm events start in the east or the central Pacific, and whether or not the incipient event is seen to propagate zonally.

While the existence of the ENSO mode depends critically on the structure of the annual average climatological state, the temporal evolution of the ENSO mode is influenced by the strong seasonality in the climatology (Fig. 8–1). By including the climatological seasonal cycle in mean state, the dynamical operator \( M \) becomes cyclostationary and the eigenmodes (now call Floquet modes) are no longer perfectly harmonic. Jin et al. (1996) showed that the spatial structure of the ENSO mode is relatively insensitive to whether the basic state includes the seasonal cycle or not (i.e., whether eigenmodes are calculated about a constant or
cyclostationary mean state). By including the seasonal cycle in the mean state, however, the time series of the ENSO mode (the leading Floquet mode) has a rich temporal structure that is consistent with observations (Jin et al. 1996; Thompson and Battisti 2000, 2001): (i) on average, warm and cold phases of ENSO occur every 3.9 years, and they tend to peak at the end of the calendar year (Fig. 8-14, top), and (ii) when the ENSO Floquet mode is forced by stochastic white noise, the spectrum of Niño-3 is peaked at interannual time scales (Fig. 8-14, bottom) and the variance in Niño-3 is peaked at the end of the calendar year; both results are consistent with observations.

Finally, Neelin and Jin (1993) explored sensitivity of the coupled modes to the strength of the coupling (i.e., the strength of the Bjerknes feedback) and to the speed of restoring ocean memory. Among their many interesting findings is that a sufficiently strong Bjerknes feedback leads to pure growth of anomalies, and there

---

Fig. 8-13. Structure of the ENSO mode in a variant of the Zebiak and Cane (1987) coupled model as revealed in the thermocline, zonal wind stress, and SST anomalies. Hovmöller plots of (top) the leading (most unstable) eigenmode of the linearized model and (middle) the same fields from the full Zebiak and Cane model (time in days increasing upward). Maps of the SST at the peak of the warm phase are shown for (left) the eigenmode and (right) the full model. [Adapted from Battisti and Hirst (1989).]
is a minimum width for ocean basins to support tropical coupled modes on interannual time scales. Zebiak (1993) showed that the Atlantic supports interannual variability that is similar in spatiotemporal structure to ENSO and that an intermediate coupled model of the Atlantic (with prescribed climatological fields) is able to reproduce the observed ENSO-like mode in the Atlantic. On the other hand, there is no ENSO-like mode in the Indian Ocean because the mean state SST is relatively uniform horizontally and there is no cold tongue due to the weak annual mean zonal wind stress (Fig. 8-1 and see section 2).

d. Further evidence in support of the ENSO mode

One critical question concerning the ENSO mode was answered toward the end of TOGA: how efficient is the western boundary at reflecting incoming Rossby waves within \( \pm 10^8 \) of the equator that are fundamental to the delayed ocean memory in the ENSO mode? Theory, results from the intermediate coupled models, and observational studies showed that on the time scales relevant to ENSO, there is plenty of reflection to support an ENSO mode (Clarke 1991; Du Penhoat and Cane 1991; Kessler 1991; Mantua and Battisti 1994).

The physics governing the warm and cold phases in the ENSO mode—and the transition from the warm phase to the cold phase by the delayed ocean dynamic memory in the ENSO mode—is remarkably similar to what is seen in observations. It is not always the case, however, that the observed warm events can be easily traced to prior ocean memory (if they could, then the observed ENSO would be more periodic). It is interesting to look back at Wyrtki’s (1975) influential study, in which he examine the relationship between wind stress and three instances of coastal warming off Peru, which also happened to be times of large-scale ocean warming (i.e., the warm phase of ENSO). Wyrtki showed that the three warm events were not due to local wind stress changes off Peru. Rather, he argued that they were due to anomalously strong easterlies in the year or two before the coastal warm event, which built up warm water in the western Pacific that was released in the form of a downwelling Kelvin wave. What he did not appreciate was that the easterly anomalies were the oceanic expression of the cold phase of the ENSO mode (all three warm events were preceded by cold events and two of the cold events were very strong). Wyrtki’s observations are in fact what is expected theoretically because the coupled ENSO mode oscillates slowly and freely between cold and warm phases. Nonetheless, taken across the entire record however, warm phases are much more likely to be followed by cold phases than vice versa (e.g., see Fig. 8-8, top), consistent with the ENSO mode being a stable mode and with greater-amplitude warm events than cold events (so the delayed ocean thermocline anomalies that transitions the warm phase to the cold phase have a greater amplitude than the thermocline anomalies that transitions the cold phase to the warm phase; see, e.g., Mantua and Battisti 1994; Kessler 2002; Kessler and McPhaden 1995).

e. Asymmetry in the amplitude of warm and cold ENSO phases

Although the essential physics of ENSO is well captured by linearized dynamical models and by linear inverse models fit to observations, the statistics of ENSO are not Gaussian. Some progress has been made in the past 20 years on identifying the source(s) of asymmetry in the amplitude of El Niños and La Niñas (i.e., positive skewness in the eastern Pacific SST anomalies). One possible source would be nonlinearity in the atmospheric response to SST anomalies, which is seen in the Coupled Model Intercomparison Project phase 5.
(CMIP5) models that tend to have unrealistically strong El Niño events such as in the Geophysical Fluid Dynamics Laboratory Climate Model, version 2.1 (GFDL CM2.1) (Choi et al. 2013; Atwood et al. 2017). Another possibility is state-dependent noise forcing—in particular, asymmetry in strength of the wind burst activity in the developing phase of warm versus cold events (e.g., Timmermann et al. 2003; Tziperman and Yu 2007; Jin et al. 2007; Chen et al. 2015; Levine et al. 2016; Levine and Jin 2017). Indeed, the zonal wind stress is positively skewed in the central Pacific (i.e., in the region of coupling), but so too is SST in the eastern Pacific [Fig. 8-15; see Choi et al. (2015) for a discussion of the dynamics associated with the asymmetry in wind stress anomalies]. However, the relationship between zonal wind stress in the central Pacific and SST anomalies in the eastern Pacific is extremely linear (Fig. 8-15; see also Battisti and Hirst 1989; Lübbecke and McPhaden 2017), the hallmark of the Bjerknes feedback. These data, and the analyses of Chiang et al. (2001) and Adames and Wallace (2017), indicate it is unlikely that the nonlinearity in ENSO resides in the atmospheric response to SST anomalies.

Fig. 8-15. Relationships between monthly Niño-3.4 index, thermocline depth in the eastern Pacific (averaged 5°S–5°N, 170°E–120°W), and zonal wind stress anomalies in the central Pacific (5°S–5°N, 160°E–150°W) for the period 1979–2017. (top) Histograms with skewness γ; (middle) scatterplots of Niño-3.4, wind stress, and (leading by 3 months) thermocline depth; (bottom) Q–Q plots of the variables against one another. The red line in the scatterplots represents the least squares fit (see text). A 5-month running mean is placed on Niño-3.4 and zonal wind stress; results are not sensitive to smoothing or the details in the region over which the data are averaged. The SST, thermocline, and zonal wind stress data were obtained from the Climate Explorer (http://climexp.knmi.nl/start.cgi) and are from ERSST v.3, Yin et al. (2011), and ERA-Interim, respectively.
Other sources of asymmetry between warm and cold phases include nonlinearity in the relationship between SST and net surface heat flux anomalies, and nonlinearity in the relationship between SST and changes in ocean circulation or mixing. Figure 8-15 shows there are strong nonlinearities between thermocline depth anomalies and SST anomalies along the equator in the eastern Pacific (see also Lübbecke and McPhaden 2017). Since studies have shown that the observed thermocline anomalies are well modeled as a linear response to zonal wind stress forcing [see section 4b(2)], this suggests that the asymmetry in ENSO rests in the SST response to dynamical changes in the ocean (e.g., Battisti and Hirst 1989; An and Jin 2004; Jin et al. 2003; Lübbecke and McPhaden 2017; Kohyama et al. 2018). Indeed, the nonlinearity in the SST response to changes in ocean dynamics is the major source of the asymmetry in the Zebiak–Cane model (Zebiak and Cane 1987) and in CMIP5 models that simulate realistic ENSO variance and skewness (Kohyama et al. 2017).

f. Irregularity of ENSO

As shown in the last section, irregularity in ENSO events is ensured because the pure, linear ENSO eigenmode exists in a cyclostationary, invariant basic state (Thompson and Battisti 2000). However, other factors likely contribute to making ENSO irregular. These fall into three general categories: (i) stochastic (uncoupled) forcing, (ii) nonlinear interaction between ENSO and other coupled atmosphere–ocean modes, and (iii) very low frequency changes in the basic climatological state that supports ENSO. Although it is difficult to quantify the relative contributions of these three factors to the irregularity in ENSO, there is strong evidence that the stochastic forcing of the nonharmonic ENSO mode is likely the dominant source of irregularity (see, e.g., Chang et al. 1997; Blanke et al. 1997; Kleeman and Moore 1997; Philander and Fedorov 2003; Kirtman et al. 2005; Lopez and Kirtman 2014).

1) STOCHASTIC FORCING

Some of the variability in ENSO must stem from stochastic forcing—variability internal to the atmosphere and/or ocean that is to be independent of coupling. Indeed, since most studies find that the ENSO mode is a damped mode of the coupled system, some sort of stochastic forcing is required to maintain ENSO variance.

Many sources of stochastic forcing have been proposed and demonstrated in an even greater number of studies using a hierarchy of coupled models, including the intermediate coupled models, full ocean models coupled to Gill-type and empirical atmosphere models (the latter are often referred to as hybrid coupled models), and coupled climate models modified to illuminate the impact of high-frequency weather on ENSO. Among the first examining the impact of stochastic forcing in the aforementioned model categories were studies by Penland and Sardeshmukh (1995), Zebiak (1989), Chang et al. (1996), and Kirtman and Shukla (2002), respectively. By the end of TOGA, more than a dozen studies had been published examining the impact of stochastic (atmospheric weather) forcing on ENSO [see the review paper by Neelin et al. (1998)], and many more papers have been published in the past two decades. Many of these studies do not attribute the stochastic forcing to specific physical phenomenon, but instead assume the forcing to be the residual wind forcing after the low-frequency variability is removed, or assume the stochastic forcing to have a certain spectrum (e.g., white in space and time).

The leading candidate for forcing is stochastic variability in the North Pacific Oscillation (Walker and Bliss 1932; Linkin and Nigam 2008), a pattern of atmospheric variability in sea level pressure that, together with the upper-level west Pacific teleconnection pattern (Wallace and Gutzler 1981; Hsu and Wallace 1985), represents meridional fluctuations of the Pacific jet stream. NPO variability arises naturally in the atmosphere and is consistent with synoptic-scale dynamical processes, including an e-folding decay time of 7–10 days (Hsu and Wallace 1985; Feldstein 2000). The relevant features of the NPO for ENSO include variations in subtropical SLP and trade wind variability that can excite the Pacific meridional mode (see section 3 and Fig. 8-5; see also Vimont et al. 2003a,b; Anderson 2003). Vimont et al. (2003b) found that an anomalous positive (negative) anomaly in the winter-averaged NPO (Fig. 8-16, top) is highly correlated with El Niño (La Niña) conditions in the following winter (Fig. 8-16, bottom). They analyzed 40 years of observations and illuminated the causal physics of this connection by performing experiments with an atmospheric general circulation model. For example, a negative NPO anomaly averaged over winter (0) is associated with anomalies in the turbulent heat fluxes that cause anomalous warming of the ocean in the subtropical North Pacific, which typically lasts from late spring to midsummer. In turn, the positive subtropical SSTs force changes in the atmospheric circulation (that last as long as the anomalies in SST) including anomalous westerlies to the south of the SST maximum that extend to the equator in the west-central Pacific. The anomalous westerlies force a train of downwelling ocean Kelvin waves that are further amplified by the Bjerknes feedback, causing El Niño conditions in the next winter (similarly, a negative NPO in
winter contributes to La Niña conditions the following winter). Vimont et al. (2003b) called this physics the seasonal footprinting mechanism (SFM) for ENSO. Chang et al. (2007) showed that the SFM accounts for nearly half of the variance and 70% of the El Niños and La Niñas in ENSO over the period 1958–2000 (Fig. 8-17; see also Park et al. 2013); similar results were found in the coupled modeling studies of Chang et al. (2007) and Alexander et al. (2010).

We note that the SFM also contributes to the pronounced seasonality of ENSO variance, because the NPO variability is maximum in the Northern Hemisphere wintertime (Fig. 8-17a) and the Bjerknes feedback is greatest in the tropical cold season July–November (Chang et al. 2007). Hence, there is a strong synergy between the seasonal timing of stochastic NPO variability and the seasonality in ENSO variance due to both the seasonality in the SFM and the seasonal cycle in the tropical climatology (Thompson and Battisti 2001).

Another suggestion for irregularity of ENSO that falls into this category is westerly wind burst activity on the equator in the west-central equatorial ocean that is commonly seen during the onsets of El Niño events (e.g., Vecchi and Harrison 2000; Kessler 2002; Kessler and McPhaden 1995; Eisenman et al. 2005; Perez et al. 2005; Batstone and Hendon 2005). Theory suggests—and modeling results and analysis of observations show—that in isolation, high-frequency wind forcing cannot impact the low-frequency variability (i.e., ENSO; Roulston and Neelin 2000; Vecchi et al. 2006; Levine and Jin 2010; Lopez and Kirtman 2014; Capotondi et al. 2018), in part because the aspect ratio of the zonal-depth profile of energy response scales as frequency $^{-3}$. However, the observed westerly wind bursts come in packets and are choreographed by the low-frequency changes in the tropical Pacific SST. As such, they can be thought of as an expression of the slow adjustment of the coupled atmosphere–ocean system (Eisenman et al. 2005; Tziperman and Yu 2007; Gebbie et al. 2007; Levine and Jin 2010) and thus are implicitly included in the ENSO eigenmode obtained from the linearized intermediate climate models (e.g., Thompson and Battisti 2000) and cannot be seen to be responsible for the irregularity in ENSO. Indeed, modeling and observational studies that include the wind burst adjustment as a multiplicative noise forcing that is a function of SST show realistic ENSO variability that is well described by the ENSO eigenmode (e.g., Eisenman et al. 2005; Jin et al. 2007; Perez et al. 2005; Capotondi et al. 2018).

On the other hand, an unknown—but likely large—fraction of the aggregate (low frequency) westerly wind burst activity along the equator in the west-central Pacific in late spring and summer is not intrinsic to the ENSO mode. Rather, it is due to the response of the atmosphere to late spring and summertime anomalies in the subtropical north-central Pacific SSTs that results from the integrated wintertime forcing associated with

---

**Fig. 8-16.** Stochastic forcing of ENSO via the SFM (updated from Vimont et al. 2003b). Shown are the leading patterns from an MCA of (top panel) SLP in boreal winter of year 0 with (bottom panel) SST one year later. The SLP pattern in (a) is nearly identical to the North Pacific Oscillation—a major pattern of variability intrinsic to the atmosphere (Walker and Bliss 1932). A random negative excursion in NPO causes positive SST anomalies in the summertime subtropical North Pacific (not shown) by way of weakened evaporation (weaker than normal trades; see Fig. 8-5). In turn, the positive off-equatorial SST anomalies cause anomalous precipitation and weakened trade winds along the equator that last for most of the boreal summer (not shown) and give rise to (bottom) El Niño conditions in the following winter via the Bjerknes feedback. Contour interval 0.2 hPa (std dev)$^{-1}$ for SLP; 0.15°C (std dev)$^{-1}$ for SST. Positive contours are solid, negative contours are dashed, and the zero line has been omitted. Color denotes correlations (contour interval 0.1). [Figure as in Vimont et al. (2003b), but updated using 1949–2018 data.]
uncoupled stochastic NPO variability (Vimont et al. 2003b) that persists for many months, rendering the SFM to be an efficient source of stochastic low-frequency wind stress forcing of ENSO.

2) NONLINEAR INTERACTION BETWEEN THE ENSO MODE AND OTHER COUPLED ATMOSPHERE–OCEAN MODES

Nonlinear dynamics may contribute to the variability of ENSO. For example, if the coupling between the atmosphere and ocean is strong enough, the ENSO mode becomes unstable (Zebiak and Cane 1987; Cane et al. 1990) and may interact with the seasonal cycle (Jin et al. 1994; Chang et al. 1995, 1996; Tziperman et al. 1994; Stuecker et al. 2015) or other unstable coupled modes to create chaotic solutions (Münnich et al. 1991; Mantua and Battisti 1995; Tziperman et al. 1995). Chaotic solutions feature variability that is qualitatively different from that observed in the modern climate (two attractor basins, bimodal statistics, spectra peaks that are not in observations, etc.; e.g., Mantua and Battisti 1995; Münnich et al. 1991; Tziperman et al. 1995; Chang et al. 1996; Timmermann 2003; Stein et al. 2014; Timmermann et al. 2018), primarily because the coupling strength used in these studies is greater than observed. There is considerable evidence from models and the analyses of observations that the major sources of ENSO irregularity are linear dynamics acting on a fixed seasonal cycle and uncoupled (stochastic) forcing, and not chaotic dynamics (see the previous section; see also Chang et al. 1997; Blanke et al. 1997; Kleeman and Moore 1997; Philander and Fedorov 2003; Kirtman et al. 2005; Giese and Ray 2011; Takahashi et al. 2011; Lopez and Kirtman 2014).

3) VERY LOW FREQUENCY CHANGES IN BASIC STATE

Intermediate models and the theory of ENSO that stems from these models strongly suggest that the spatial and temporal structure of the ENSO mode is sensitive to the structure of the climatological mean state and to the spatial and temporal distribution of the uncoupled stochastic forcing [see, e.g., Hirst (1986, 1988), Zebiak and Cane (1987), and Battisti (1988) and the review studies by Neelin et al. (1998) and Chang et al. (2006)]. Hence, it is to be expected that very low frequency changes in the mean state will affect the statistics and spatial structure of ENSO (e.g., Fedorov and Philander 2000). Moon et al. (2007) argued that the low-frequency modulation of ENSO variance in the Center for Ocean–Land–Atmosphere (COLA) climate model was due to low-frequency changes in the stratification of the near-equatorial ocean that stemmed from wind stress forcing in the subtropical southern Pacific Ocean. Matei et al. (2008) demonstrated that SST anomalies placed in the southern subtropical Pacific of the ECHAM5–Max Planck Institute Ocean Model (MPI-OM) coupled model eventually make their way to the equator and cause mean state changes that affect the variance in ENSO. Gu and Philander (1997) suggested that multidecadal variability in the tropical Pacific may come about by way of feedbacks between the tropics and extratropics. Specifically, they proposed that ENSO-driven anomalies cause oppositely signed anomalies in SST in the subtropics (as observed) that are subducted and returned to the equator many years later by the shallow, wind-driven mean ocean overturning circulation, but analyses of observations do not support this.
hypothesis (see, e.g., N. Schneider et al. 1999; Hazeleger et al. 2001).

Potential sources of low-frequency variations in the tropical Pacific mean state that may drive multidecadal changes in ENSO variability include low-frequency variations from other regions and basins, such as the PDO (see section 6b) and low-frequency variability from the Indian or Atlantic Ocean basins. For the Pacific, if the PDO features changes in wind stress that extend deep into the subtropics in the central Pacific, the PDO would force multidecadal changes in the strength of the subtropical ocean overturning circulation that would affect the temperature of the water injected into the undercurrent and/or the mean thermocline depth along the equator (Hazeleger et al. 2001), thereby causing long-lived changes in the background state that ENSO experiences. Although ocean current data are available for less than two PDO “cycles,” there is a remarkable correspondence between the multidecadal PDO and Zhang and McPhaden’s (2006) 48-yr record of the strength of the subtropical overturning circulation, and between the subtropical ocean overturning circulation and the low-pass-filtered SST in the equatorial eastern Pacific 1 or 2 years later. With only 1.5 multidecadal PDO cycles in these data, however, it is not possible to tell whether these low-frequency mean state changes affect interannual ENSO variability.

In addition to the PDO, low-frequency variability in the tropical Atlantic and Indian Oceans has been implicated as a cause of low-frequency modulation of ENSO variability (see review by Cai et al. 2019). The Atlantic and Indian Oceans may impact the background mean state upon which ENSO evolves through adjustment of the large-scale Walker circulation in the tropics, atmospheric equatorial wave propagation, or changes in the midlatitude storm tracks (Dong et al. 2006; Zhang et al. 2007; Luo et al. 2012; McGregor et al. 2014). Resulting changes in thermocline depth and SST can subsequently affect characteristics of ENSO variability itself, such as amplitude, frequency, and spatial structure (Dong et al. 2006; Levine and Jin 2017). While some modeling efforts support these mechanisms, the instrumental record is short, making it difficult to draw definitive conclusions about the role of forcing from other basins.

Some studies with climate models illustrate how nonlinear dynamics associated with ENSO might rectify to cause very low frequency changes in the tropical seasonal cycle that may, in turn, affect the statistics and stability of ENSO (e.g., Timmermann 2003). Other studies have suggested that stochastic noise—independent of ENSO—may drive mean state changes that in turn affect ENSO statistics (e.g., Burgman et al. 2008). In a cautionary tale, however, Wittenberg (2009) analyzed a very long unforced simulation of a version of the GFDL coupled model (CM2.1) and showed the model featured long (more than 40 years) periods in which the variance of ENSO differed by more than a factor of 4. However, the difference in ENSO variance in these multidecadal epochs was due to differences in the history of the uncoupled weather forcing that the ENSO mode experienced during those epochs (Wittenberg et al. 2014) and not due to differences in the mean state; indeed, Atwood et al. (2017) showed that the mean state changes were likely due to the rectified effects of ENSO and worked to oppose the changes in ENSO variance.

It is unclear whether nature can display a fourfold change in variance due to noise alone (i.e., without mean state changes): as is true of nearly all coupled general circulation models, there are notable qualitative differences between the ENSOs simulated in the GFDL model and those observed (see section 4g). However, there is substantial evidence that the modern climate features a damped ENSO mode that is maintained by stochastic uncoupled forcing [see sections 4c and 4f(1)], and so is inevitable that the spatial and temporal evolution of ENSO must vary significantly even with no change in the mean state. Indeed, studies with stochastically forced LIMs show that the variance of ENSO should wax and wane on multidecadal time scales—without changes in the stability of the ENSO mode and without mode–mode interactions—simply due to the stochastic nature of the forcing. These models readily feature flavors of ENSO (e.g., central Pacific ENSOs or eastern Pacific ENSOs) even though they have no mean state changes—simply due to the stochastic nature of forcing and/or the superposition of linear coupled modes (e.g., Penland and Sardeshmukh 1995; Newman et al. 2011b; Vimont et al. 2014; Capotondi and Sardeshmukh 2015).

g. Response to forcing

Ray and Giese (2012) and Yang and Giese (2013) examined SST from an ocean reanalysis product covering the period 1871–2008 and concluded there is no evidence in the historical record of changes in the characteristics (strength, frequency, duration, location) of El Niño or La Niña events. Upon examining the results from the scores of twentieth- and twenty-first-century simulations from the high-end climate models used in the last two IPCC assessments (the CMIP3 and CMIP5 models), Bellenger et al. (2014) also concluded “there is low confidence that ENSO has changed in the twentieth century.” and they go on to say that “there is no conclusive evidence on how ENSO properties may change in the future.” Indeed, CMIP5 models do not
agree on the sign of the change in ENSO amplitude or frequency (Cai et al. 2015).

This conclusion, however, does not imply ENSO will not change due to anthropogenic forcing. With few exceptions, the comprehensive (CMIP5) climate models have a difficult time simulating realistic ENSO cycles (see Bellenger et al. 2014; Capotondi et al. 2015a; Yun et al. 2016; Takahashi et al. 2011; Guilyardi et al. 2016, and references therein). As expected from theory, biases in ENSO are correlated with biases in the simulated climatology (Neale et al. 2008; Guilyardi et al. 2009; Zhu et al. 2017). For example, El Niño in both the CMIP3 and CMIP5 climate models feature maximum SST anomalies that are too far west of those observed, consistent with a prominent westward bias in the cold tongue location (see Capotondi et al. 2015b, and references therein); only 40% of the models capture the observed structure of ENSO, as measured by the first two EOFs of SST (Yun et al. 2016); the amplitude of the simulated shortwave feedbacks in the modeled ENSO are only half of that in observations (Bellenger et al. 2014); the simulated Bjerknes feedback is typically 20%–50% weaker than that in observations, with less than 20% of the models within 25% of the observed feedback strength (Bellenger et al. 2014); models have warm and cold phases of ENSO that are too symmetric compared to that observed (Yang and Giese 2013); and less than half of models have realistic ENSO variance (as measured by the variance in standard indices of ENSO; Yun et al. 2016). Will models with more realistic mean states and realistic ENSOs in their twentieth-century simulations simulate large changes in ENSO that are consistent across the models? It remains to be seen.

The alternative hypothesis—that ENSO is insensitive to the spatial structure and amplitude of the anthropogenic forcing—is unlikely. Projections show anthropogenic forcing will very likely increase the stratification in the upper tropical ocean significantly (Timmermann et al. 1999; Collins 2000) and affect the asymmetry in the climatological mean state of the upper ocean—both of which should greatly affect ENSO. Indeed, it is not uncommon for a CMIP model to simulate a large change in ENSO by the end of this century (e.g., Yeh et al. 2009; Cai et al. 2015), but models do not agree on how ENSO will change; for example, increased ocean stratification is thought to have greatly increased ENSO variance and frequency in the ECHAM4–OPYC3 and Hadley Centre Coupled Model (HadCM2), coupled models, but to have caused a large decrease in ENSO variance in the GFDL Earth System Model (GFDL-ESM2M; Timmermann et al. 1999; Collins 2000; Kohyama et al. 2017). In many CMIP5 models certain aspects of the mean state changes due to greenhouse gas or orbital forcing, in insolation, greatly increase the variance of ENSO (e.g., Clausius–Clapeyron feedbacks) while other aspects tend to reduce the variance of ENSO (e.g., a weakening of the Walker circulation; see DiNezio et al. 2012; Roberts et al. 2014). The net impact on ENSO depends on a delicate balance of many processes (Zebiak and Cane 1987; Battisti 1988), and thus it is to be expected that biases in the mean states of the models greatly affect the projected changes in ENSO due to anthropogenic forcing.

CMIP5 models have also been used to simulate the response of ENSO to volcanic eruptions and to changes in insolation due to orbital changes. We do not review these studies here because of the limited number of simulations.

h. ENSO summary

ENSO appears to be a damped eigenmode of the linearized coupled atmosphere–ocean system in the tropical Pacific that features atmosphere–ocean coupling in the central-eastern equatorial Pacific (the Bjerknes feedback) and delayed adiabatic ocean memory. The spatial and temporal structure of the ENSO mode is strongly influenced by the annual mean climatological basic state, while the spectra and the phasing of ENSO to the calendar year is largely determined by the seasonal cycle in the basic state and seasonality in the stochastic forcing (mainly through the seasonality of the SFO). Modeling and theoretical results suggest that variability in ENSO is governed more by these processes than by chaotic variability or interaction between two or more modes of coupled variability.

Although the observational record is short, it is largely consistent with this theory for ENSO cycles and the irregularity of ENSO. There is strong evidence that ocean memory effectively terminates most El Niño events and gives rise to subsequent cold conditions (La Niñas), but a substantial number of El Niño events do not appear to be triggered by the ocean memory from a previous La Niña event. This result is consistent with an ENSO mode that is strongly damped, and with the observations (Okumura and Deser 2010) that warm events

---

9 In particular, most models suffer the so-called double ITCZ problem: a strong ITCZ south of the equator that extends from the central Pacific eastward, instead of very dry conditions (observations show a weak ITCZ south of the equator that is limited to February–April; see Fig. 8-1). As such, the annual averaged southerly wind stress in the eastern Pacific is too weak. As well, the seasonal cycle in SST and zonal wind stress along the equator are also too weak in the CMIP5 models compared to observations (Chen and Jin 2017)—a bias that is expected with a double ITCZ.
are almost always followed the next year by cold events, but not vice versa. Nonlinearities in the dynamics (yet to be fully understood; see section 8) cause warm (El Niño) phases to be more extreme than cold (La Niña) phases, rendering greater amplitude in the restoring ocean memory that terminates warm phases (and ensures a following cold phase) than terminates cold phases (Mantua and Battisti 1994). Hence, it is not surprising that uncoupled stochastic wind forcing is seen to play a greater role in the initiation of warm events than cold events. The different flavors of ENSO are consistent with stochastic forcing of a damped ENSO mode that results from a rich blend of processes that govern SST in space and time without changes in climatological basic state or an external forcing (see, e.g., Zebiak and Cane 1987; Battisti 1988; Penland and Sardeshmukh 1995; Giese and Ray 2011; Newman et al. 2011b; Vimont et al. 2014; Capotondi and Sardeshmukh 2015; Lai et al. 2015).

i. The tropical Atlantic zonal mode

We briefly mentioned in section 4c that the tropical Atlantic features interannual variability that in many ways is analogous to ENSO in the tropical Pacific, with irregular cycles of warm and cold SST events focused along the equator that are accompanied by anomalies in the zonal wind stress and thermocline depth. This mode of variability is now called the tropical Atlantic “zonal mode.” Indeed, Zebiak (1993) built an intermediate atmosphere–ocean model (see section 4c) of the tropical Atlantic that successfully simulated the observed zonal tropical Atlantic zonal mode. He argued that the physics of the zonal mode was similar to that of ENSO: an oscillating mode that resulted from the tension between a positive Bjerknes feedback and a restoring ocean memory. Compared to the Pacific ENSO, he found that the Atlantic zonal mode features a stronger coupling (Bjerknes feedback) that is offset by the smaller basin size (hence, a shorter ocean memory)—resulting in a coupled atmosphere–ocean mode with a period of about 4 years that was asymptotically stable. Zebiak also noted that there were some distinct differences between the Atlantic zonal mode and ENSO, including a more equatorially trapped SST anomaly and anomalies that peaked in the middle of the calendar (ENSO peaks at the end of the calendar year).

Subsequent studies affirmed Zebiak’s conclusions that the tropical Atlantic zonal mode has physics that are akin to the physics operating in ENSO (Chang et al. 2006; Keenlyside and Latif 2007; Burls et al. 2012; Lübbecke and McPhaden 2017) and have provided further insight to the differences in physics of the Atlantic zonal mode and ENSO mode. For example, unlike in the Pacific, in the Atlantic there appears to be a significant seasonal dependence on both the strength of the Bjerknes feedback (Keenlyside and Latif 2007; Lübbecke and McPhaden 2017) and the restoring ocean memory (Burls et al. 2012).

5. The waxing and waning of the Indian Ocean dipole

Saji et al. (1999) defined an Indian Ocean dipole mode index (DMI) as the SST difference averaged over the western (10°S–10°N, 50°–70°E) and eastern (10°S–0°N, 90°–110°E) Indian Ocean. These regions were not chosen on physical grounds, but on their being the two (oppositely phased) centers of action of the second-leading EOF of Indian Ocean SST anomalies, explaining 12% of the total SST variance. Regression of dipole index on zonal wind shows zonal wind anomalies along the equator in the east-central Indian Ocean. Hence, Saji et al. hypothesized a coupled atmosphere–ocean phenomenon with dynamics akin to ENSO in the Pacific Ocean.

Many studies have extended Saji et al.’s (1999) analysis of the observations to include upper-level heat content, while other studies have examined output from global climate models to see whether the Indian Ocean supports a true coupled atmosphere–ocean zonal dipole-like mode that includes active upper-ocean dynamics. A challenge for each of these investigations was to isolate variability that is intrinsic to the Indian Ocean from that forced by ENSO, as the latter is responsible for the lion’s share of variability in the Indian Ocean [furthermore, as noted by Saji et al., there is a statistically significant correlation between ENSO and DMI, and Meyers et al. (2007) found half of the events in the DMI time series were associated with ENSO].

Collectively, the observational studies do not support a zonal dipole mode intrinsic to the Indian Ocean. A straightforward correlation of SST in the two boxes that comprise the DMI index show that SST is uncorrelated in the months of July–November—the very months the dipole mode is thought to exist (Zhao and Nigam 2015); this result is not sensitive to the period sampled, from 1900 to 2017, or to whether or not ENSO is removed from the data. Furthermore, the variance in the DMI is overwhelmingly due to SST variability in the eastern Indian Ocean. Observational studies show westerly wind stress anomalies along the equator in the central Indian Ocean are correlated with positive anomalies in upper-ocean heat content (a proxy for thermocline depth) in the eastern half of the Indian Ocean, and that the spatial pattern is consistent with the basin adjustment to wind-forcing (Chang et al. 2006; Zhao and
However, the SST anomalies (concurrent and delayed) associated with the wind stress anomalies are located south of Indonesia and do not extend to the equator or west of about 90°E; that is, the data do not support the hypothesis that thermocline variations affect SST beyond the coast of Indonesia or along the equator (Saji et al. 1999; Chang et al. 2006; Zhao and Nigam 2015). SST anomalies in the eastern basin off Indonesia are due to changes in upwelling, mixing, and surface fluxes associated with in situ wind anomalies (Hendon 2003; Shinoda et al. 2004) as well as remotely forced changes in thermocline depth (e.g., Murtugudde and Busalacchi 1999).

Consistent with the analyses of observations, models that support realistic patterns of SST variability in the Indian Ocean generally do not support the hypothesis that there is a coupled zonal dipole mode of variability. On the other hand, a few models do feature a coupled atmosphere–ocean mode that is akin to the ENSO mode in the Pacific. In these models, easterly zonal wind stress anomalies along the equator in the east-center basin in boreal midsummer give rise to thermocline variations and/or upwelling in the eastern basin that creates cold conditions south of Indonesia in late summer to early fall (as in the observations) that decay by the end of the calendar year. However, unlike in the observations, most of the models also feature SST anomalies along the equator that reach westward to 165°E (e.g., Iizuka et al. 2000; Behera et al. 2006; Yang et al. 2015), which is symptomatic of a climatological thermocline that is much shallower in these models than in observations (Saji et al. 2006) and is strong evidence that the Bjerknes feedback is much more important for the coupled variability simulated in these models than in observations (Schott et al. 2009). Not surprisingly, models that feature a too-strong Bjerknes feedback also feature SST variability that is much greater than that observed (Yang and Giese 2013; Liu et al. 2014; Créaté et al. 2017; Cai and Cowan 2013); on average, the amplitude of a typical excursion in the DMI index in the CMIP3 and CMIP5 models is twice that observed (Cai and Cowan 2013; Li et al. 2015).10

6. Coupled atmosphere–ocean interaction in the midlatitudes of the Northern Hemisphere

In the Northern Hemisphere, orography and the substantial large-scale diabatic heating gradients associated with the land–sea distribution give rise to localized storm tracks and large-scale stationary waves with greatest amplitude in winter. In turn, the inhomogeneity in the climatological flow organizes the variability on seasonal and longer time scales into (mostly) distinct spatial patterns of variability. The two leading patterns are called the North Atlantic Oscillation (NAO; Fig. 8-18; Walker 1924) and the Pacific–North American pattern (PNA; Fig. 8-19; Wallace and Gutzler 1981).11 The NAO signifies variability in the strength and a meridional shift of the Atlantic storm track (Wettstein and Wallace 2010); as such, the source of energy for the NAO is the baroclinity associated with the climatological temperature gradients in the lower troposphere. The PNA signifies extension and contraction of the climatological jet in the Pacific. The PNA is a favored mode of variability because it is the most rapidly growing mode of the zonally asymmetric climatological barotropic flow in the troposphere (Simmons et al. 1983); it derives its energy from section 4f(1).

---

10 Biases in the strength of the Bjerknes feedback have been attributed to biases in the climatology of the Indian Ocean in the CMIP5 models (Cai and Cowan 2013; Li et al. 2015).

11 The North Pacific Oscillation is another leading pattern that plays a central role in energizing ENSO; see section 4f(1).
from the kinetic energy of the stationary waves and is sensitive to changes in the intensity and distribution of convection in the western Pacific (Simmons et al. 1983; Sardeshmukh and Hoskins 1988). Transient eddy fluxes also contribute to the maintenance of the PNA and NAO (Franzke et al. 2011).

The NAO and PNA are the leading patterns of atmospheric variability on seasonal-to-interannual time scales. They are intrinsic to the atmosphere and its underlying geometry (their existence does not depend on interactions with the ocean) and their power spectra are nearly white on time scales longer than about 10 days (Wunsch 1999; Feldstein 2000; Wills et al. 2019a). However, these patterns give rise to changes in the ocean that, in turn, contribute to the low-frequency (seasonal and longer time scales) power in the NAO and PNA, as described below.

**a. Seasonal to interannual variability**

Russ Davis published a remarkable study in Davis (1976). Using 28 years of data, he calculated the leading EOFs of monthly SLP and SST variability in the North Pacific and examined their autolag and covariance structures (the leading EOF of SLP depicts variations in the Aleutian low and is now recognized as the surface expression of the PNA; Trenberth and Hurrell 1994). Davis showed that the leading EOFs of SST and SLP are related and that the present SLP pattern is related to the SST pattern several months in the future, but not vice versa. He deduced that, to leading order, stochastic (uncoupled, nearly white noise) atmospheric variability drives the variability in the SST on time scales of months to at least seasons. At the same time, Hasselmann (1976) proposed a simple model for the response of the slow components of the climate system to variability in fast components, and in the following year Frankignoul and Hasselmann (1977) showed that Hasselmann’s model was consistent with Davis’s calculations and interpretation. Specifically, nearly white stochastic atmospheric forcing of the ocean gives rise to a red spectrum in SST because of the high thermal inertia of the upper ocean. Later studies showed that in the midlatitudes, the forcing of the ocean by the atmosphere is largely accomplished through the latter’s modulation of the turbulent fluxes at the interface of the two fluids (Frankignoul 1985; Cayan 1992a,b), with secondary contributions from ocean advection due to wind-driven anomalies in Ekman transport (Frankignoul 1985; Deser et al. 2010).

Barsugli and Battisti (1998) extended Hasselmann’s model to include thermodynamic coupling between the
atmosphere and ocean (i.e., allowing the slow variability of SSTs to feed back onto anomalies in atmospheric temperature and surface heat fluxes) and showed that coupling enhances the persistence and variance in the intrinsic atmospheric patterns and amplifies the response of the ocean on seasonal and longer time scales (see Fig. 8-20). For example, compared to the uncoupled atmospheric variability (i.e., with fixed climatological SST), the interannual variance in the intrinsic atmospheric patterns of variability (PNA, NAO, etc.) more than doubles when the atmosphere and ocean are coupled (Fig. 8-20, top).

The time scale \( \tau \) beyond which coupling matters to the midlatitude atmosphere is about 4 months and is primarily determined by (i) the strength of the coupling between the atmosphere and ocean, (ii) how well the SST anomaly produced by the atmospheric pattern of variability projects back onto that pattern, (iii) the radiative efficiency of the atmosphere, and (iv) the thermal inertia of the ocean that participates. Essentially, the ocean acts as a passive amplifier of the otherwise white spectrum of intrinsic uncoupled atmospheric variability (weather)—a phenomenon called “reduced thermal damping” (Barsugli and Battisti 1998). On seasonal to interannual time scales, reduced thermal damping typically doubles the variance of the patterns of atmospheric variability (such as the NAO and PNA). Since most of the power in atmospheric motions is on subseasonal time scales, however, the change in total atmospheric variance due to thermal coupling with the upper ocean is only a few percent. A consequence of coupling is that on time scales longer than \( \tau \), coupled variability that is intrinsic to the atmosphere features vanishingly small surface fluxes, with the atmosphere forcing the SST anomalies (Fig. 8-20, bottom).

Many modeling and observational studies over the past 40 years have confirmed that seasonal to interannual variability intrinsic to the midlatitude atmosphere and ocean is well described by Hasselmann’s theoretical model, extended by Barsugli and Battisti (1998) to include thermodynamic coupling [see, e.g., Wallace et al. (1990, 1992), Battisti et al. (1995), Bladé (1997), Bhatt et al. (1998), and the review paper by Kushnir et al. (2002b)]. That the leading NAO and PNA patterns of atmospheric variability give rise to the SST anomalies is evident from the close correspondence between the patterns of surface turbulent flux anomalies and the SST anomalies that accompany these atmospheric patterns, shown in Figs. 8-18 and 8-19, respectively. Not shown in Figs. 8-18 and 8-19 are the changes in radiation due to cloud feedbacks, which amplify the flux anomalies in the southern lobe of the tripole in the Atlantic (Yuan et al. 2016; Brown et al. 2016; Bellomo et al. 2016). Other leading patterns of intrinsic atmospheric variability, such as the NPO that is central to the forcing of ENSO [see section 4f(1)], force SST anomalies in a similar fashion.

The Barsugli and Battisti (1998) framework for midlatitude atmosphere–ocean variability explained rather surprising results reported by Rodwell et al. (1999), Mehta et al. (2000), and others (Bretherton and Battisti 2000). Each of these studies summarized results from an ensemble of simulations using an atmosphere model forced by the observed SST variability with slightly different atmospheric initial conditions, focusing on the ability of the models to reproduce the observed winter [December–February (DJF)] averaged NAO. Although

---

12 Not surprisingly, the SST patterns in Figs. 8-18 and 8-19 are also the leading patterns of SST variability in the North Atlantic and North Pacific.
individual ensemble members poorly reproduce the observed NAO [correlation $-0.2 (-0.15, +0.42)$], the ensemble-averaged NAO is highly correlated with the observed NAO (correlation $+0.43$ to $+0.75$), suggesting there may be high skill in forecasting atmospheric variability. However, the amplitude of the ensemble averaged NAO is greatly reduced, and the air–sea fluxes are the reverse of that observed. Bretherton and Battisti (2000) showed that the Barsugli and Battisti (1998) model reproduced all of these findings, and that they were consistent with stochastic intrinsic atmospheric variability forcing a passive coupled, thermodynamic ocean with forecast skill limited two one or two seasons.

Two final points: first, since the atmosphere is most active in wintertime, the impact of stochastic atmospheric forcing is greatest in winter, when the ocean mixed layer is deep. The memory of the impact of the atmosphere on the ocean is extended from the end of one winter to the beginning of the next because upper-ocean temperature anomalies laid down in winter are sheltered from the atmosphere in the following summer when a shallow mixed layer reforms in response to the climatological summertime weakening of wind mixing and increased insolation. The mixed layer deepens again in fall, exposing the relic temperature anomalies from the end of the previous winter to the surface. This phenomenon has become known as “ocean reemergence” (Namias and Born 1970, 1974; Alexander and Deser 1995) and acts to redden the spectrum of SST. Second, ENSO contributes to the variance shown in Fig. 8-19 by way of a wintertime atmospheric teleconnection from the tropical Pacific to the North Pacific [see also section 6a], which has become known as the atmospheric bridge (although the teleconnection pattern is somewhat modified from the PNA pattern intrinsic to the atmosphere; Alexander 1992; Lau and Nath 1996).

b. Decadal to multidecadal variability

The NAO and PNA, by their stochastic nature, also have spectral power at decadal and multidecadal time scales and are thus integral to our understanding of variability on these time scales. On interannual and longer time scales, the impact of stochastic variations in the NAO and PNA on the ocean includes the local response described in section 6a and the dynamical adjustment of the ocean by way of Rossby waves (Frankignoul and Müller 1979; Saravanan and McWilliams 1998; Neelin and Weng 1999; Saravanan et al. 2000; Cessi 2000; Marshall et al. 2001; Chhak et al. 2006, 2009). Prominent patterns of variability on time scales of many years to many decades include an unnamed pattern of decadal variability in the subpolar Atlantic and the leading patterns of decadal and multidecadal variability in the midlatitude Pacific and Atlantic Oceans: the PDO and AMO, respectively. These decadal and multidecadal phenomena share a common anatomy. Stochastic forcing by the leading patterns of atmospheric variability (the PNA for the PDO, the NAO for the AMO) that create (i) heat flux anomalies that drive local SST anomalies (see Figs. 8-18 and 8-19) and (ii) surface wind stress anomalies that drive Ekman current anomalies and delayed ocean gyre current anomalies [and, for the AMO, delayed anomalies in the Atlantic meridional overturning circulation (AMOC)] that further affect SST. As such, the variance in the PDO (AMO) is predominantly decadal (multidecadal). Below we summarize what is known about these patterns of variability.

1) Decadal variability in subpolar Atlantic

Curry and McCartney (2001) examined 49 years of hydrographic, surface flux, and wind forcing data and concluded that the subpolar gyre in the North Atlantic exhibited decadal-scale variability due to the integrated, stochastic wind stress and diabatic forcing associated with the NAO. A period with an integrated positive NAO is accompanied by a cooling of the subpolar gyre and a spinup of the cyclonic subpolar gyre circulation (mainly due to pressure gradients associated with buoyancy changes); similarly, negative NAO causes a warming and a spindown of the subpolar gyre. The delayed baroclinic adjustment by baroclinic Rossby waves accounts for the power at the decadal time scale. Their findings were confirmed and refined by Deshayes and Frankignoul (2008), who forced an ocean model with the history of wind and buoyancy forcing over the North Atlantic (1948–2003) and quantified the relative roles of Ekman transport, geostrophic adjustment, and buoyancy forcing in the spatiotemporal evolution of the circulation in the subpolar gyre (see also Eden and Willebrand 2001; Eden and Jung 2001).

Deshayes and Frankignoul (2008) showed that the time-integrated NAO also governs the strength of the deep western boundary current along the western boundary of the Labrador Sea, and hence affects the strength of the AMOC that is central to the longer-term mode of climate variability known as the Atlantic multidecadal oscillation [see section 6b(3)]. The link between NAO-induced buoyancy fluctuations and the strength of AMOC is also seen in fully coupled climate models and in ocean models forced by idealized NAO variability (e.g., Yeager and Danabasoglu 2014; Lohmann et al. 2009; Medhaug et al. 2012; Mecking et al. 2014). Although in most models the NAO gives rise to damped variability in the subpolar gyre, it has been argued that observations support a positive feedback between the
subpolar ocean variability and the NAO (Czaja and Frankignoul 1999, 2002; Gastineau and Frankignoul 2015), which results in a preferred frequency of multidecadal variability in at least one model (Menary et al. 2015).

2) THE PACIFIC DECADAL OSCILLATION

Summary: Many model studies have arrived at a common description of the anatomy of the PDO (see, e.g., Schneider et al. 2002; Zhang and Delworth 2015; Newman et al. 2016; Wills et al. 2019a) that is supported by analyses of observations (e.g., Frankignoul and Hasselmann 1977; Miller et al. 1997, 1998; Qiu 2003; Kwon et al. 2010; Wills et al. 2018). These studies conclude that the PDO is a phenomenon that is intrinsic to the midlatitudes and is due to stochastic atmospheric forcing of the ocean in the form of variability in the Aleutian low—the prominent lobe of the PNA that resides over the North Pacific Ocean. The stochastic buoyancy and wind stress forcing gives rise to (i) a pattern of SST anomalies that reflects the pattern of surface heat flux anomalies associated with PNA and (ii) anomalies in the ocean gyres that lag the low-frequency changes in the wind by 3–6 years and cause subsequent anomalies in SST associated with meridional displacements in the Kuroshio–Oyashio Extension (KOE; see Fig. 8-21). The SST anomalies in the KOE region force the atmosphere, but it is not yet clear whether the response further affects the PDO.

Mantua et al. (1997) coined the moniker “Pacific decadal oscillation” to describe the time series of the leading EOF of SST variability north of 20°N in the Pacific. When this time series was regressed onto SST for the entire Pacific, the resulting patterns and amplitudes of SST and SLP anomalies are very similar to those shown in Fig. 8-19, which were obtained by regressions against the leading principal component (PC) of SLP in the North Pacific and associated with variability in the PNA. As was noted by Mantua et al. (1997), the midlatitude SLP and SST patterns associated with the PDO index are qualitatively similar to those associated with ENSO, as evidenced by regressing SLP and SST against high-pass (6-yr high-pass filtered) indices of ENSO (Zhang et al. 1997). This is to be expected because the PNA is a pattern of variability that is intrinsic to the midlatitude atmosphere. Changes in the location of tropical convection (say, associated with ENSO) are preferentially expressed in the northern midlatitude atmosphere by anomalies in a PNA-like pattern (Simmons et al. 1983), particularly when ENSO SST anomalies extend into the central Pacific (Newman et al. 2016). One might expect, then, that the PDO would look like the sum of the contributions from the ubiquitous uncoupled stochastic variability in the PNA (see section 6a) and from variations in the Aleutian low (PNA-like) that are forced directly by ENSO. In turn, the spectrum of the PDO would look like the weighted sum of the red spectrum shown in Fig. 8-20 and the spectrum of ENSO (scaled by the projection of ENSO onto the PNA).

That ENSO can contribute to the PDO is supported by analyses of the Pacific SST data. Vimont (2005) calculated the three leading EOFs of SST data that were filtered to exclude all variability on time scales greater than a decade; these leading EOFs have clear physical interpretations associated with the interannual ENSO cycle. He then showed that the decadal-scale pan-Pacific SST variability can be reconstructed using EOFs calculated from the high-frequency data. This result implies that pattern of the PDO can result from a simple time averaging of the residual of subdecadal variability.

But how much of the variability in the PDO stems from uncoupled stochastic variability and how much stems from ENSO? The correlation between the PDO and ENSO indices hovers near 0.35, indicating only a small fraction of the variability in the PDO is associated with ENSO (see also Yeh and Kirtman 2005). Indeed, the time series of the PDO features decadal scale variability and a spectrum that is similar to that expected from uncoupled stochastic forcing, rather than from ENSO (which has power concentrated between 3 and 7 years). Another indication that the PDO is primarily a midlatitude phenomenon stems from the different amplitudes of the nominal midlatitude PNA SLP pressure anomalies associated with the PDO (2 hPa) and with ENSO (1 hPa) (Wills et al. 2018, supporting information).
Studies of simulations using CMIP5 models show the midlatitude portion of the PDO to be a robust pattern of decadal variability across the models (e.g., Newman et al. 2016; Farneti 2017; Wills et al. 2019a). Most models feature PDO variability in the midlatitudes that is comparable in amplitude to that observed (see, e.g., Fig. 3 of Farneti 2017). However, not all models extend the PDO footprint into the tropical and South Pacific (Newman et al. 2016)—which, combined with the observational evidence above, further suggests the PDO is predominantly a midlatitude Northern Hemisphere phenomenon that is due to stochastic PNA forcing and redounded by a thermodynamic coupling with a passive underlying ocean (reduced thermal damping; section 6a). Wills et al. (2018) used low-frequency component analysis (LFCA)\(^\text{13}\) to extract from observations the pattern of PDO variability that has the greatest ratio of low frequency to total variance, LFC-PDO (Fig. 8-21). The time series of the LFC-PDO pattern is highly correlated with the original PDO index (the first PC of North Pacific SST, \(r = 0.86\)). LFC-PDO is only weakly related to variability in the tropical Pacific (correlation \(-0.15\); see also Yeh and Kirtman 2005). Hence, the LFC-PDO SST pattern shown in Fig. 8-21 retains all of the midlatitude features in the original PDO, but it de-emphasizes the variability in the tropical central Pacific. LFCA finds patterns that maximize the ratio of low frequency to total variance, such that LFC-PDO captures a greater fraction of the low-frequency SST variance in the Pacific than the traditional PDO. As a result, LFC-PDO has much more persistence than the PDO, and at long lead times (>1.5 years) is a better predictor of the PDO than PDO itself (Wills et al. 2018).

The circulation of the upper ocean is also affected by stochastic PNA forcing, through anomalies in Ekman transport and Ekman pumping. On seasonal to interannual time scales Ekman transport augments (albeit a minor contribution) the surface turbulent flux anomalies that give rise to the large-scale SST anomalies (Frankignoul 1985; Seager et al. 2000; Pierce et al. 2001; Kushnir et al. 2002a). On longer time scales, Ekman pumping generates westward-propagating Rossby waves and an adjustment of the midlatitude ocean gyres that is delayed from the wind stress forcing by several (3–6) years (Frankignoul and Hasselmann 1977) and seen in both observations and models (Frankignoul and Hasselmann 1977; Chelton and Schlax 1996; Miller et al. 1998; Deser et al. 1999; Wills et al. 2019a). For example, Miller et al. (1997) examined thermocline data and showed that stochastic variations in the Aleutian low cause westward-propagating Rossby waves that cross the entire Pacific in 5–7 years and cause anomalies in SST in the Kuroshio–Oyashio Extension region that force the atmosphere (as evidenced by the negative correlation between SST and turbulent heat flux from the atmosphere to the ocean in the KOE region; see also Pierce et al. 2001; Qiu 2003; Qiu and Chen 2005; Taguchi et al. 2007; Kwon et al. 2010; Sasaki and Schneider 2011, and references therein).

So is the PDO a dynamically coupled mode of variability, as was originally proposed by Latif and Barnett (1994, 1996)? The lion’s share of the PNA and SST variability is certainly due to passive thermodynamic coupling as described by the extended Hasselmann model (Frankignoul and Hasselmann 1977; Barsugli and Battisti 1998). However, it is unresolved whether the delayed SST response in the KOE region due to the ocean dynamical adjustment further influences the PDO. Some modeling studies report the delayed SST anomalies in the KOE region force the atmosphere in a way that provides a negative feedback onto the Aleutian low (PNA), giving rise to multidecadal-scale oscillations (Zhang and Delworth 2015), as was first hypothesized by Latif and Barnett (1996). And as noted in several recent papers, the atmospheric resolution in the CMIP5 models may be inadequate for capturing the response to the small-spatial-scale SST anomalies in the KOE and Gulf Stream exit regions (e.g., Smirnov et al. 2015). However, most modeling and observational studies show weak positive, ambiguous, or no atmospheric response to the delayed SST anomalies in the KOE region (e.g., Saravanan et al. 2000; Schneider et al. 2002). At best, studies report a response that is sensitive to the background climatology and strongly dependent on the time over which averaging is done [see Revelard et al. (2018) for an excellent discussion and summary of these studies].

3) THE ATLANTIC MULTIDEcadAL OSCILLATION

Summary: An extensive literature on the multidecadal AMO phenomenon has accumulated over the past few decades. In this section, we summarize the dynamics of the AMO as gleaned from these observational and modeling studies and place in footnotes the myriad references in support of each mechanistic step of the AMO. Although these studies portray a mostly consistent

\(^{13}\)LFCA weights the linear combination of EOFs (calculated using unfiltered data) to identify patterns of variability that maximize the ratio \(r\) of low-frequency variance to total variance in a dataset using covariance on all time scales. Although LFCA finds patterns that maximum \(r\), it also allows for rapid transitions in the patterns and does not dilute important phase information. Hence, it is particularly efficient for illuminating the physics of the PDO and AMO. See Wills et al. (2018, 2019b) for details.
description of the physics of the AMO (see, e.g., Wills et al. 2019b, and references therein).\textsuperscript{14} we caution the reader that this picture relies heavily on the analysis of CMIP3 and CMIP5 climate models and features coupling to the AMOC. Although analyses of observations support many of the mechanistic steps in this picture of the AMO, observations are not adequate to confirm the role of AMOC. Hence, the story below remains a hypothesis.

The physics of the AMO as seen in climate models is revealed by low-frequency component analysis and depicted schematically in Fig. 8-22. The driver for the AMO is thought to be stochastic variability in the NAO, or in models with less realistic variability, ocean internal variability of the AMOC. As discussed in section 6a and shown in Fig. 8-18, the positive phase of the NAO is associated with increased wind speeds and cold air advection over the western portion of the subpolar gyre, coincident with one of the regions of deep water formation in the North Atlantic (more so in models than in the very limited observations). The immediate effect is

\textsuperscript{14} Note that Zhang and Wang (2013) and Ba et al. (2014) report less agreement across the models with regard to AMO physics, but this is due in part because they define the AMO index as the basin-wide average SST rather than an SST index focused in the subpolar gyre.
to reduce the SST (Fig. 8-18) and reduce the buoyancy of the upper ocean, which enhances the AMOC (see Buckley et al. 2014, and references therein). Because of the great inertia of the ocean, the AMOC response is an integral of the NAO forcing and lags the NAO by several years.\(^1\)\(^5\) Hence, a decade or so in which the average NAO is stronger than normal causes an increase in the AMOC strength (Fig. 8-22a).

An increase in AMOC brings heat from the subtropics to the subpolar region, which overcompensates for the immediate cooling due to the positive NAO forcing and causes a warm anomaly to develop in the subpolar gyre (Fig. 8-22b).\(^1\)\(^6\) Unlike on seasonal to interannual time scales, these longer-lived positive SST anomalies are associated with anomalous heat flux from the ocean to the atmosphere—unequivocal evidence of the active and preeminent role of ocean circulation in the AMO (cf. Clement et al. 2015).\(^1\)\(^7\) Indeed, the positive SST anomaly in the subpolar gyre is sustained by the ocean heat flux convergence associated with the anomalously strong AMOC and lasts many years. The atmosphere responds to this heat flux by creating a low pressure anomaly (Hoskins and Karoly 1981) that is essentially orthogonal to the NAO (cf. the colored contours in Figs. 8-22a and 8-22b). The atmospheric anomaly generated by the AMOC-driven surface heat flux anomalies is associated with warm air advection in the northeastern Atlantic and reduced trade winds in the subtropical eastern Atlantic, both of which cause an anomalous heat flux into the ocean that spread the warm anomaly from the western subpolar gyre to the eastern and subtropical Atlantic (Chiang and Friedman 2012; cloud feedbacks also amplify in the subtropical response).\(^1\)\(^8\) The buoyancy gain in the subpolar gyre scales roughly with the SST anomaly, and hence the AMOC anomaly peaks ~2–3 years before the SST peaks. In all but one CMIP5 model, the atmospheric (wind stress) response to the SST anomalies does not project onto the NAO, and so there is no direct feedback of the atmospheric response onto the AMOC. Hence, the AMOC continues to spindown due to buoyancy loss in the subpolar gyre (Fig. 8-22c) and therefore the positive SST anomalies can no longer be sustained, and the warm phase of the AMO decays (Wills et al. 2019b).

The same physics works when starting from a sufficiently long enough period of time in which the average of the stochastic NAO variability produces a negative AMO anomaly. Hence, the AMO can be thought of as a first-order Markov process in which the stochastic forcing of the AMOC is provided by the NAO (e.g., Mecking et al. 2014; Wills et al. 2019b, and references therein). Unlike on seasonal to interannual time scales, however, the time scale \(\lambda\) of the AMO is determined by the scaling of heat flux convergence associated with a unit of AMOC variability, the efficiency of heat exchange between the atmosphere and ocean in the western subpolar gyre, and the dynamical inertia of the AMOC [which is poorly understood; see the discussion in Buckley and Marshall (2016)]. From observations, \(\lambda\) is estimated as the decorrelation time of the AMO index, which is \(~7\) years or so, and so the time scale of AMO variability is many decades (Figs. 8-22d and 8-22e; Wills et al. 2019b).\(^1\)\(^9\)

The basic anatomy of the AMO described above and summarized in Fig. 8-22 is robust across the CMIP5 climate models and is largely consistent with the available observational records of SLP, SST, and surface heat fluxes that span more than a half century (Wills et al. 2019b).\(^2\)\(^0\) The amplitude of the AMO varies in the models, but the scaling of the amplitude of the AMOC changes to the amplitude of the SST changes is remarkably similar across the models (Medhaug and Furevik 2011; Yan et al. 2018; Kim et al. 2018). Furthermore, the evolution of the SST pattern in the AMO (Fig. 8-22) is also qualitatively similar across models: maximum SST variability is in the far northern Atlantic (usually collocated with the location of deep convection in each model), and most models feature the eastward and southward extension of the warming that immediately follows the peak warming in the subpolar Atlantic.

\(^{19}\) We note that a different mechanism for multidecadal variability in the Atlantic is at work in the GFDL CM2.1 model, which features a strong 20-yr spectral peak (MacMartin et al. 2013). The time scale of the oscillation in this model is likely set by the ocean basic state and independent of coupling to the atmosphere. The T85 version of the Community Earth System Model (CESM) also features an AMO with a strong multidecadal spectral peak, but the AMO variability in that model is unrealistic (Kwon and Frankignoule 2012; Frankignoule et al. 2013).

\(^{20}\) The link between AMOC and SST variability on decadal time scales was also reported in early vintage climate models (Delworth and Greatbatch 2000; Delworth et al. 1993; Latif et al. 2004), but the AMOC-related SST and SLP patterns in these models did not show good agreement with observations.
The observational record of AMOC is too short to directly confirm its role in the AMO, but the limited observations show a high correlation of AMOC strength and the structure of the upper-ocean heat content in the subpolar North Atlantic, and (as a proxy for AMOC strength) the latter data support the role of AMOC variability in the AMO phenomenon as described above. Similarly, uncoupled ocean-ice models forced by the observed NAO forcing reproduce the observed decadal- and multidecadal-scale changes in the subpolar North Atlantic, and do so by changing the strength of AMOC (Danabasoglu et al. 2016). Finally, although not seen in the multimodel mean or in individual CMIP5 models, analyses of observations suggest subpolar SST anomalies may project onto the NAO (see Frankignoul et al. 2017, and references therein), which, if strong enough, would tend to produce a preferred time scale of variability. However, the relationship is fragile to the choice of season used for both the SST and circulation anomalies, to whether and how trends and the forced signals are removed from the data (see, e.g., Booth et al. 2012a), and to the methodology used to relate SST and atmospheric circulation.

Is the AMO described herein a coupled mode of variability (as defined in section 1)? On the one hand, the atmospheric phenomenon (the NAO) that forces the AMOC anomalies is a phenomenon that is intrinsic to the atmosphere, but the SST anomalies that result from the AMOC changes are different from those expected from a simple slab ocean response to NAO forcing. On the other hand, the collective modeling and observational evidence suggests the atmospheric response to the AMOC-generated SST anomalies does not directly project back onto the AMOC (Hodson et al. 2010; Sun et al. 2015; Wills et al. 2019b).

7. Seasonal to interannual predictability

The presence of slowly evolving variations at the surface of Earth (e.g., sea surface temperature, soil moisture, snow cover, and sea ice) is the scientific basis for predicting seasonal-to-interannual climate variability (Shukla 1998). As an example, consider the case of El Niño conditions; that is, warm SST anomalies in the tropical Pacific. These positive SST anomalies enhance the flux of sensible heat and moisture from the ocean into the atmosphere. Locally, this enhanced flux modifies the atmospheric boundary layer leading to large-scale shifts in tropical rainfall patterns and an associated release of latent heat in the midtroposphere (top panel of Fig. 8-7). The shift in the midtropospheric latent heating is primarily balanced by changes in atmosphere circulation, which ultimately results in remote teleconnections throughout the globe (see reviews by Trenberth et al. 1998; Yeh et al. 2018). One well-documented remote teleconnection, for instance, is the enhanced winter rainfall over the southeastern United States during El Niño conditions in the tropical Pacific (e.g., Lyon and Barnston 2005; Kirtman et al. 2017). Given the fact that the El Niño conditions can persist for many months, this suggests that the remote teleconnections will also persist for several months. Moreover, as the discussion in section 4 suggests, the El Niño conditions can be predicted several months in advance, again indicating that there is potential to predict the remote teleconnections several months in advance. While ENSO is the dominant source of seasonal-to-interannual climate predictability, similar processes can be detected in the tropical Atlantic and Indian Oceans, and there are ongoing efforts in the PREFACE project to mine this predictive skill (see https://preface.w.uib.no). Although less well understood, extratropical SST can also be a source of predictability [see Kirtman et al. (2013) for additional discussion].

In terms of the discussion of seasonal-to-interannual prediction here, much of our focus is on the prediction of slowly evolving SST anomalies, particularly associated with ENSO. Nevertheless, it is important to note that there are other important boundary conditions (e.g., soil moisture; see Delworth and Manabe 1989; Koster et al. 2000; Koster and Suarez 2003; Koster et al. 2004; Seneviratne et al. 2006) and phenomena (e.g., meridional modes; see section 3) that evolve slowly and are potentially sources of predictability. However, considerable progress has been made in understanding and predicting ENSO, and this section seeks to summarize some of this progress with two foci in mind: (i) predicting the onset and life cycle of individual ENSO events and (ii) predicting the associated remote teleconnections.

a. Predicting the onset and life cycle of ENSO

The theoretical and observational understanding of ENSO matured in the 1980s (see section 4). This understanding clearly identified that equatorial western Pacific thermocline anomalies were a potential source of equatorial eastern Pacific SSTA predictability (McCreary 1983; Suarez and Schopf 1988; Battisti 1988; Battisti and Hirst 1989; Schopf and Suarez 1990). This basic understanding motivated Cane et al. (1986) to use the celebrated Cane–Zebiak model (Zebiak and Cane 1987) to make the first dynamical prediction of the onset of an El Niño event in 1986. At the time, the forecast was controversial, but ultimately it proved useful in supporting

---

21 Zhang (2008); Mahajan et al. (2011); Huang et al. (2012).
the implementation of the TOGA Program on Prediction (TPOP). One of the outcomes of TPOP was the establishment of the Experimental Long-Lead Forecast Bulletin (ELLFB; cola.gmu.edu/ellfb/past.html) that was first published by NOAA in 1992 and was transferred to the Center for Ocean–Land–Atmosphere Studies (COLA) in 1998, where it continued through 2011. The ELLFB provided a forum for testing and evaluating predictions from a wide range of models: empirical models (e.g., Penland and Magorian 1993; Drosdowsky 1994; Barnston et al. 1994; Keppenne and Ghil 1992; Tangang et al. 1997; Knaff and Landsea 1997), simplified dynamical models (e.g., Chen et al. 1995; Kleeman et al. 1995), combined dynamical and empirical hybrid models (Balbasteda et al. 1994; Barnett et al. 1993), and complex coupled general circulation models (GCMs; Leetmaa and Ji 1989; Kirtman et al. 1997). The ELLFB also laid out some ground rules for facilitating the evaluation of prediction systems, including the documentation of methods used and the reproducibility and forecast quality assessments.

Early on, empirical-based forecasts were arguably more skillful predictions than forecasts made with dynamical climate models. This was particularly true for broad-brush indices such as Niño-3.4. More recently, the dynamical systems have become somewhat more skillful than empirical prediction systems. Dynamical systems, of course, have the added advantage of predicting the entire state of the climate system, whereas statistical prediction systems typically predict single indices and select climate variables. Certainly, statistically predicted SST can be used to force an atmospheric GCM in a so-called two-tiered system. This has the potential advantage of removing the systematic error before the SSTs are communicated to the atmosphere, but has the disadvantage of inconsistent air–sea fluxes of heat and momentum.

Most of the growth in ENSO prediction has focused on the development of coupled GCMs. Delecluse et al. (1998) provide a detailed summary of this development during the TOGA decade, and Latif et al. (1998) summarize the ENSO forecast skill of these systems. During the 2000s, a number of operational meteorological centers and research groups started producing real-time ENSO predictions with coupled GCMs (E. K. Schneider et al. 1999; DeWitt 2005; Zhang et al. 2007; Merryfield et al. 2013; Kirtman 2003; Kirtman and Min 2009; Wang et al. 2002; Ji et al. 1994; Ji 1996; Ji et al. 1998; Stockdale et al. 1998a; Rosati et al. 1997; Molteni et al. 2011; Jungclaus et al. 2013; MacLachlan et al. 2015; Weisheimer et al. 2009; Takaya et al. 2017; Voldoire et al. 2013, among others). For example, the Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER) project (Palmer et al. 2004; Hagedorn et al. 2005; Doblas-Reyes et al. 2005) systematically assessed the quality of forecasts made using the coupled GCMs at several operational European centers with one particularly powerful result—namely, that a multimodel ensemble probabilistic forecast is more skillful and more reliable than any single model. The international CLIVAR Climate-System Historical Forecast Project (CHFP; Kirtman and Pirani 2009) and the North American Multi-Model Ensemble (NMME; Kirtman et al. 2014) project capitalized on this multimodel result to emphasize that forecast ensemble spread, essentially a forecast of the forecast error, should line up with the actual forecast error. This notion is exemplified in Fig. 8-23 (taken from Tompkins et al. 2017), which shows the Niño-3.4

![Fig. 8-23. Model skill in predicting Niño-3.4 (1982–2009) as a function of lead time for prediction. For the individual models the RMSE (black) is calculated based on the difference of the individual model ensemble mean from the observational estimates. For the individual models, the ensemble standard deviation (or deviation from the ensemble mean) is shown in red. The RMSE for the multimodel ensemble (MME) is based on the average of the ensemble means from each model (i.e., the models are weighted equally). [From Tompkins et al. (2017).]
root-mean-square errors and forecast ensemble spread from each of the nine CHFP models, and in the multimodel ensemble. Clearly, the individual models struggle to capture the spread–error relationship, whereas the spread and error line up well in the multimodel ensemble.

One of the biggest challenges in seasonal-to-interannual prediction is overconfident forecasts. Put simply, with our current generation of models, any single model can lock onto a particular phase (e.g., forecasting ENSO warming too early). The multimodel approach, by including models that have complementary behavior, is a pragmatic way to increase model spread and effectively reduce overconfidence. This is exactly what has led better consistency between forecast spread and forecast error in Fig. 8-23. This reduced overconfidence is often quantified by assessing reliability, and, in fact, the multimodel prediction system is generally more reliable than any single model (see Figs. 10 and 11 in Kirtman et al. 2014). The improved reliability is due to the fact that models have structural uncertainty (e.g., different parameterized physical processes) and multimodel methodologies can, in part, quantify this uncertainty. And, with sufficiently long periods of retrospective forecasts, the models could ultimately be weighted based on past performance, although this remains a challenge with current retrospective forecast samples. While multimodel is a pragmatic strategy, other strategies are emerging that may prove more effective—in particular, perturbed physics ensembles, which is being done at the U.K. Met Office, and stochastic physics approaches (primarily advanced by ECMWF; e.g., Berner et al. 2008).

The multimodel (or model diversity) approach for quantifying forecast uncertainty is also the basis for NMME project that began issuing monthly and seasonal forecasts in 2011 and was officially declared a NOAA operational system in 2016. The NMME team examined the effect of increased ensemble size versus model diversity. To make this comparison, they formulate a “mini-NMME” that uses multiple models but with the exact same ensemble size (24) as NOAA’s nominal operational Climate Forecast System, version 2 (CFSv2; Saha et al. 2014). Table 8-1 shows the Brier skill score (BSS; provided by H. van den Dool 2015, personal communication) for retrospective Niño-3.4 forecasts for 1982–2010 (all month starts) for the CFSv2, the mini-NMME, and (with over 100 members in the ensemble) the full NMME. For the Niño-3.4 SST anomaly (and tropical SST anomalies in general, not shown), model diversity counts for more skill than the increased ensemble size. The improvement in forecasts of near-normal Niño-3.4 conditions (i.e., the middle category of a tercile probabilistic forecast) due to model diversity is particularly notable. In terms of the Northern Hemisphere extratropical SSTA, ensemble size and model diversity contribute equally to the improvement in forecast skill (not shown).

The so-called spring prediction barrier begs some discussion. ENSO SSTAs tend to peak in the boreal winter season and are either in transition or are near normal during the boreal spring. Despite improvement in models, better observing systems, and the implementation of multimodel systems, the forecast quality of boreal spring tropical Pacific SSTA is notoriously low. One explanation is simply that the SSTAs are small in amplitude and not particularly large scale. More recent work by Larson and Kirtman (2017) argues that the coupled system is particularly sensitive to unpredictable stochastic forcing in the boreal spring.

b. Predicting the remote teleconnection associated with ENSO

While it is clear that ENSO is the main driver of seasonal-to-interannual predictability, current operational systems like NMME use state-of-the-art coupled GCMs that include the interactions among all components of the

---

22 The reader is referred to Kirtman et al. (2014) for a full description of the NMME system.

23 A BSS of 0.0 is equivalent to a climatological forecast, and a BSS of 0.33 is considered “good” skill.
climate system (i.e., atmosphere, ocean, land surface, sea ice) and even changes in atmospheric composition (e.g., CO₂). In principle, these predictions are capturing all the potential sources of predictability. Here we show briefly some results from the NMME project (additional details and a more comprehensive view are available at www.cpc.ncep.noaa.gov/products/NMME/). First, we show the deterministic global SSTA forecast quality (correlation coefficient) for the individual models comprising the NMME and for the multimodel ensemble in Fig. 8-24. The correlation coefficient is calculated for retrospective forecasts initialized each July of 1982–2010 and verifying the following December–February for each year. Clearly, the models have different levels of skill and it is possible to identify the model that has the best skill (even comparable to the multimodel skill) with this particular metric, forecast initial month, and lead time, whereas the multimodel is always comparable to the best model.

The probabilistic skill in the forecast of DJF SST starting from July conditions is shown in Fig. 8-25; each ensemble member for each model has been equally weighted, yielding about 100 members in the ensemble average forecast. The probabilistic skill is measured using the ranked probability skill score (RPSS), which effectively measures the root-mean-square difference between the predicted probability and the observations. RPSS has the advantage that any increase can be compared to the deterministic skill of the best model.

24 RPSS is a probabilistic forecast skill metric [see Weigel et al. (2007) for details]. The RPSS evaluates the quality of the forecast probabilistically, using tercile-based categories and the equal-odds climatology forecasts as the reference forecast. A good rule of thumb is that an RPSS of 0.3 corresponds to a deterministic correlation of 0.7.
directly related to economic value in a cost–loss decision model sense. Larger values of RPSS indicate larger skill, and thus, more potential economic value.

Ultimately, the purpose of seasonal-to-interannual prediction is to produce forecasts that are of societal value. This implies, as a first step, forecasts of climate variables that are of societal interest. Indeed, one of the advantages of using comprehensive coupled GCMs is that they produce predictions for the state of the entire climate system. Here we show as an example (in a similar format as Figs. 8-24 and 8-25) the multimodel deterministic and probabilistic forecast skill for precipitation (Fig. 8-26). The rainfall forecast skill is notable in the tropical Pacific and in some of the well-known regions of robust teleconnections (e.g., southeast United States, tropical South America, and South Africa).

The results presented here suggest considerable optimism for the utility of seasonal-to-interannual climate predictions. There are, however, many outstanding questions, and notable forecast failures pointing to areas that require improved understanding, modeling capability, and prediction systems. For example, the multimodel ENSO forecast for the winter of 2014/15 was far too confident in predicting El Niño conditions that failed to occur (Larson and Kirtman 2015), suggesting the need to reassess the metrics used to measure forecast uncertainty, and the need for a better understanding of ENSO predictable dynamics. The warm event of the winter of 2015/16 was well predicted well in advance, but the rainfall teleconnections that were forecasted for the west coast of the United States failed to materialize—again, pointing to either a poor representation of forecast uncertainty or biases that cut across all the models.

While there are successes to champion, quantifying the limits of predictability remains an open question, and the processes that ultimately limit predictability are not particularly well understood. For example, we do not have a clear understanding of whether predictability is fundamentally limited by uncoupled weather noise or internal nonlinear dynamics (i.e., chaos). At present, however, the primary limitation on the seasonal-to-interannual (in particular, ENSO) forecasts skill in the current coupled GCMs is likely systematic model error, rather than weather or the efficacy of its representation in the GCMs (Stan and Kirtman 2008), and capturing the details of the SSTA evolution beyond the classical Niño indices is also quite limited (Infanti and Kirtman 2016). Focused research and development will be required if we are to address these issues and further extend the capability for seasonal-to-interannual climate prediction to benefit society.

**Fig. 8-25.** RPSS for the grand multimodel ensemble for SST forecasts. Forecasts were initialized every July (1982–2010) and verified the following DJF.

**Fig. 8-26.** (left) Precipitation anomaly correlation ($\times 100$) and (right) RPSS for the grand multimodel ensemble precipitation forecasts initialized each July 1982–2010 (i.e., 28 years of retrospective forecasts) and verifying in the following DJF. The grand multimodel ensemble includes over 100 ensemble members for seven different models. The verification dataset is CMAP (Xie and Arkin 1997a).
8. **On the horizon, under the rug, and over the rainbow**

We briefly summarize in this section advances in coupled atmosphere–ocean variability that are likely to be realized in the near future (say, in the next decade), thorny problems that we wish would go away (but will not), and barriers to progress that are likely to exist for decades.

*a. On the horizon*

1) **Increased understanding of the atmospheric response to midlatitude ocean variability**

Recent work has shown that at ocean frontal scales and mesoscales, the atmosphere and ocean are tightly coupled (e.g., Chelton et al. 2004; Chelton and Xie 2010; Li and Carbone 2012; Smirnov et al. 2015). Furthermore, there is growing evidence that the large-scale atmospheric circulation and storm tracks may be affected by midlatitude SST anomalies that have sharp gradients (e.g., Minobe et al. 2008; Sampe et al. 2010; Kwon et al. 2011; Frankignoul et al. 2011; Booth et al. 2012b; O’Reilly and Czaja 2015; Ma et al. 2015, 2017; Kuwano-Yoshida and Minobe 2017; Putrasahan et al. 2017; Foussard et al. 2019)—gradients that are on a scale that is too small to be resolved by the current generation of atmosphere models used in climate modeling. Hence, modest increases in resolution in both the atmosphere and ocean that are achievable in the near term will illuminate whether the atmospheric response to PDO and AMO SST anomalies is such that it gives feedback onto the relevant atmospheric structures (the PNA and NAO, respectively) that are the main drivers of the low-frequency midlatitude variability, and thus it should help inform on whether decadal variability in the midlatitudes is due to dynamically coupled interactions between the atmosphere and ocean in the midlatitudes.

2) **Increased predictability on seasonal and interannual time scales**

In general, skillful predictions of ENSO are limited to 6–9 months in advance: considering all initial condition months and a robust retrospective forecast period, the correlation between the observed and forecast Niño-3.4 is typically above 0.6 for lead times of 6–9 months in state-of-the-art prediction systems. There are, however, several examples of “forecasts of opportunity,” where skillful forecasts extend to 18–24 months in advance. Work needs to be done to identify a priori situations whereby forecasts will be skillful well in advance, but this is something that is on the near-term horizon. There are other sources of multiyear predictability that will likely be mined in the near-term. For example, the persistence of the PDO is a source of predictability that may well extend out to ~5 years but has not been fully utilized to date—in part because models have notoriously underestimated the persistence of the PDO. However, models are improving in this regard and we expect that the considerable persistence of the PDO (and AMO) will be a useful source of predictability.

3) **Improved model representation of processes central to ENSO**

As described earlier, progress is already being made in identifying possible causes of extreme ENSO events, especially the positive skewness of SST in the eastern equatorial Pacific. Current research has found that this skewness can be caused by nonlinearity (section 4e) or multiplicative noise (sections 4e and 4f). We expect that advances in understanding and modeling of tropical convection, as well as improved resolution of ocean processes, will likely lead to more realistic representation of ENSO in models in the near future. A caveat to this expectation, though, is that improved representation of these processes may require more realistic simulation of the tropical Pacific mean state (see below), which could preclude advances in our understanding in the short term.

4) **Improved understanding of the regional response to aerosols**

While our understanding of the full response of the tropical Pacific—and ENSO—to anthropogenic forcing will likely elude us until model mean states are better represented, it is likely that we will be able to assess the direct (e.g., the partial derivative) response to regional aerosol forcing. This is especially important in assessing the spatial structure and impacts of aerosol forcing in the Atlantic (Ting et al. 2009; Chang et al. 2011; Chiang et al. 2013; Tandon and Kushner 2015), and in understanding heat content variations in the Pacific and globally (Clement and DiNezio 2014). We emphasize that while we expect increased understanding of the regional footprint of external forcing and its projection onto known modes of ocean–atmosphere variability and their impacts, the role of external forcing in altering the characteristics of ocean–atmosphere variability will be elusive and will require improvements in model mean states and physical process representation.

*b. Under the rug*

**Better historical datasets to evaluate seasonal-to-interannual predictability**

One of the most important aspects in developing seasonal-to-decadal prediction systems is to perform
retrospective forecasts. These retrospective forecasts serve many purposes, including identification of systematic model error and forecast quality assessments, and they are aided by better state estimates provided by coupled and uncoupled reanalyses products. For seasonal-to-interannual prediction, the retrospective forecast periods are on the order of 30–35 years. Given that warm ENSO events occur approximately every 2–7 years, the number of degrees of freedom is on the order of, at best, 15 and, at worst, 5. Moreover, the climate is nonstationary during the retrospective forecast period, and there are sources of external forcing (e.g., volcanic and anthropogenic aerosols) that may affect the forecasts. Indeed, we do not have a clear sense of whether the twentieth-century forecast skill is fully indicative of the future skill, and whether very low frequency modes of climate variability exist that can moderate seasonal-to-interannual forecast skill. All of these factors suggest that our retrospective forecast skill may not be indicative of future skill. Indeed, some of the recent experience with ENSO forecasts suggests that we have not adequately quantified forecast skill. Altogether, there is need to develop longer historical records of observational datasets that are suitable for forecast initialization and evaluation. These issues are even more challenging on decadal time scales—if we seek to develop a forecast of the PDO and the AMO. Indeed, there are several additional datasets that are suitable for forecast initialization and evaluation. These issues are even more challenging on decadal time scales—if we seek to develop a forecast of the PDO and the AMO. Indeed, there are several efforts at seeking to advance initialized decadal predictions (e.g., Smith et al. 2013), which at this stage appear most promising in the North Atlantic.

c. Over the rainbow

1) REDUCTION OF CLIMATE MODEL BIASES

Reduction of biases in the climate simulated by the climate models and increased resolution for explicit simulation of critical processes are particularly important issues for improving the simulation of coupled atmosphere–ocean variability (in particular, ENSO), for achieving the full potential for climate prediction on seasonal to multidecadal time scales, and for assessing the impact of climate change on the patterns of internal variability (e.g., on ENSO, the AMO, the PDO). The present generation of climate models exhibit certain biases in their climatology that greatly limit progress in understanding the patterns of internal variability associated with atmosphere–ocean interaction, as well as how these phenomena may change due to forcing. For example, biases in the simulated climatological seasonal cycle of the tropical atmosphere and ocean are certainly a major reason for the gross discrepancies between the ENSOs simulated by the CMIP5 models and the observed ENSO. Model biases in the climatological orientation of the Atlantic storm track cause biases in the climatological circulation of the wind-driven ocean currents. Biases in the position of the Gulf Stream, in the location of deep convection, and in the strength of the climatological ocean overturning circulation in the North Atlantic are sure to compromise the efficacy of the climate models for simulating the internal decadal and multidecadal variability in the North Atlantic.

Many of the biases in climate models are long-standing: for example, the so-called “double ITCZ” in the tropical Pacific has been a feature of most climate models for time immemorial. Increased resolution afforded by increased computational power has in some cases helped to reduce biases in the climate models (e.g., in the southwest-to-northeast tilt of the Atlantic storm track). In particular, it is very likely that increasing the resolution of atmosphere models will render a more realistic response of the atmosphere to midlatitude SST anomalies, and increased ocean model resolution may lead to more realistic simulations of the midlatitude SST anomalies associated with ocean fronts and mesoscale eddies. For example, Siqueira and Kirtman (2016) identify multiyear variability that is only present with a well-resolved Gulf Stream, and Putrasahan et al. (2016) show that eddies and a resolved Agulhas current are essential for capturing ENSO oceanic teleconnections in the Southern Ocean. In other cases, however, it will be quite some time before sufficient computational power exists to explicitly resolve processes that are thought to be important for some patterns of climate variability (e.g., eddies and convection in the ocean, and clouds in the atmosphere).

2) DETECTION AND ATTRIBUTION OF CHANGES IN INTERNAL MODES OF VARIABILITY

It is clear that characteristics of various modes of ocean–atmosphere interaction are strongly influenced by processes that operate through the mean state; as a result, future changes to the mean state are very likely to impact variability associated with these modes. However, large biases in the mean state, and in existing models’ representation of physical processes that we know are important, will preclude advances in predicting specific changes to the characteristics (e.g., the intensity, frequency, spatial structure, or teleconnections) of ocean–atmosphere variability in the near future. Further, the long time scales and known event-to-event variability of ocean–atmosphere variations implies that predicted changes may not be detectable or attributable in the foreseeable future. Toward this end, we see an important role for understanding past changes (in the paleo-record) in both mean climate and variability.
9. Our reflections

The summary of progress over the last 100 years in understanding coupled ocean–atmosphere interactions is, for all practical purposes, an entire history of the field. This opportunity to reflect back on the major advances in our understanding, then, provides an opportunity to consider some lessons learned.

Perhaps the simplest lesson learned from the history of progress in understanding coupled ocean–atmosphere variability is that remarkable phenomena can emerge when multiple systems interact and constrain each other through even a small number of simple physical processes. While the feedback loop described in Bjerknes (1969) is an excellent—although perhaps not the first—example of such recognition for the ocean and atmosphere, we use the framework outlined in Hirst (1986) as a more nuanced example. There, two dynamic systems (the ocean and atmosphere) are coupled using a very simple set of physical processes (linearly parameterized heating and surface drag). In the absence of those coupling parameters, each system contains its own set of modes that describe the spatiotemporal variability that system can produce. However, when the two systems are coupled, a new set of interactions emerge that permit growth of new structures that owe their existence to the coupling itself. Importantly, the coupling constrains the spatial and temporal phasing between the modes in each system, which, depending on the nature of the phasing, can lead to enhanced growth or decay of specific structures of variability. The critical role of coupling in providing phase relationships that constrain the joint evolution of the two systems is a transferable lesson that must be considered as our field advances toward more complete representation of the multiple systems that compose our climate.

A second lesson learned is that advances in understanding benefit from coordinated efforts to improve observational networks, modeling frameworks, and theoretical understanding. Those advances in understanding are valuable, even if the broader impacts of that understanding are not known a priori. The TOGA program provides a clear example of progress in understanding ENSO variability that included design and implementation of a novel observation system, development of new numerical modeling frameworks, and evolution of our theoretical understanding. This work led to the first numerical model predictions of ENSO and seasonal forecasting. These advances show that a well-coordinated effort at advancing understanding can lead to unanticipated benefits beyond the original scope. Programs that enable that holistic approach, such as U.S. CLIVAR, are novel in our community and should be recognized for providing a voice for the scientific community in advising research priorities in our field.

Finally, while immense progress has been made in the last 100 years, we recognize the excitement of things to come. New advances in observational methods, modeling capabilities, and process-level understanding holds promise for amazing new insights in the field of coupled ocean–atmosphere variability from the next generation of scientists. And if the reader has managed to stay with us this far, it is worth a moment to take a deep breath and reflect on the beauty and excitement that this field of coupled ocean and atmosphere offers. Standing along the Pacific shore in Ecuador, looking west, the gray of the sea and sky looks unassuming; but underneath the dull ocean surface, beyond the seemingly endless stratus, and within the mind of the curious scientist, yet lies immense potential.

Acknowledgments. The authors thank George Kiladis, Ping Chang, and an anonymous reviewer—as well as Joseph Barsugli, Robb Jnglin Wills, Michael Alexander, Antonietta Capotondi, and Luke Parsons—for critical and helpful comments on versions of this manuscript. DSB was supported by a grant from the Tamaki Foundation. BPK was supported by NSF OCE1419569, OCE1559151, and NOAA NA15OAR4320064, NA18OAR4310293. DJV was supported by NSF CLD 1463970. NCEP Reanalysis and CMAP Precipitation data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, from their website at www.esrl.noaa.gov/psd.

APPENDIX

List of Acronyms

- AMM: Atlantic meridional mode
- AMO: Atlantic multidecadal oscillation
- AMOC: Atlantic meridional overturning circulation
- ASO: August, September, October
- CESG: Cross-equatorial SST gradient
- CHFP: CLIVAR Climate-System Historical Forecast Project
- CLIVAR: Climate Variability and Predictability Program
- CMIP: Coupled Model Intercomparison Project
- CTI: Cold tongue index
- DEMETER: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction
- DJF: December, January, February
- ELLFB: Experimental Long-Lead Forecast Bulletin
- ENSO: El Niño–Southern Oscillation
- EOF: Empirical orthogonal function

Unauthenticated | Downloaded 04/11/22 12:58 AM UTC
REFERENCES


Gyfford, B., P. Bracaninot, F.-F. Sin, S. T. Kim, M. Kolasinski, T. Li, and I. Musat, 2009: Atmospheric feedbacks during...


La中有中文文字


——, P. Chang, R. Saravanam, R. Montuoro, H. Nakamura, D. Wu, X. Lin, and L. Wu, 2017: Importance of resolving Kuroshio front and eddy influence in simulating the North Pacific storm...


