stones for ENSO modeling. Gill (1980) showed that a single vertical mode linear model could capture the major features of the tropical atmosphere's response to the anomalous heating associated with variations in tropical SSTs. Rasmusson and Carpenter (1982) synthesized the incomplete and often perplexing observational fragments into a coherent picture of the evolution of the "canonical" ENSO warm event. Though undoubtedly a simplification, this invaluable distillation provided the first specific target for models to emulate.

18.3 Our understanding of the mechanisms of ENSO: Pt II

The Bjerknes–Wyrtki theory did not explain the perpetual turnabout from warm to cold states and back again. The explanation turns on some special features of the equatorial ocean dynamics that govern the variations in the upper tropical ocean over time scales of a few years or so. In this context, it is adequate to regard the ocean as a two-layer system: an active upper layer separated by the thermocline from a deep abyssal layer of greater density. Motions in the lower layer are quite slow, and may be disregarded relative to the active layer above. Making the useful assumption that linear physics govern the evolution of vertically integrated characteristics of this upper layer such as thermocline displacements and upper layer transports, the governing equations reduce to the shallow water equations of tidal theory.

The total response may be analyzed into a sum of free and forced waves. For the time and space scales relevant to El Niño, only two types of wave motions matter: long Rossby waves and equatorial Kelvin waves. The latter are strongly trapped to the equator and owe their existence to the vanishing of the Coriolis force there. Except within a few degrees of the equator, the geostrophic balance between Coriolis and pressure gradient forces dominates the dynamics of the ocean. This balance is characteristic of Rossby waves and strongly constrains their propagation speeds and their amplitude in response to wind driving. The Kelvin wave propagates eastward and is the fastest of the low-frequency ocean motions. It can cross the Pacific in less than three months whereas the fastest Rossby wave is three times slower.

Long Rossby waves propagate energy westward, while Kelvin waves travel to the east. When a Kelvin wave hits the eastern boundary, its reflection is made up of an infinite sum of Rossby waves, which collectively act to extend the equatorial wave guide up and down the eastern boundary to high latitudes. However, the faster Rossby waves at low latitudes carry much of the mass and energy brought east by the Kelvin waves back toward the west. The reflection process at the west is more efficient: all of the mass flux which the Rossby waves carry into the boundary is collected by boundary currents and brought equatorward, where it is returned eastward in the form of Kelvin waves.

The particular properties of equatorial waves mentioned above are essential to the El Niño phenomenon. Only at low latitudes can low-frequency waves cross the ocean in times matched to the seasonal variations in the winds. A given wind change generates a stronger response at the equator than at higher latitudes, and equatorial waves are less susceptible to the destructive influences of friction and mean currents. Finally, as will be seen, the asymmetries in the waves and their reflections are essential to the ENSO cycle. While the explanation for the perpetual oscillations of the ENSO cycle is inherent in the physics just described, it emerged only after the development of the numerical models discussed in Sec. 18.4.

An early version (Cane and Zebiak, 1985; Cane et al, 1986) emphasized the recharging of the equatorial "reservoir" of warm water as a necessary precondition for the initiation of a warm event. On the basis of his analysis of sea level data, Wyrtki (1985) developed a very similar hypothesis. The aftermath of a warm event leaves the thermocline along the equator shallower than normal (i.e., equatorial heat content is low and SST is cold; this is the "La Niña" phase). Over the next few years the equatorial warm water reservoir is gradually refilled. Once there is enough warm water in the equatorial band, the rapid (for the ocean) equatorial Kelvin waves allowed by linear equatorial ocean dynamics can move enough of the warm water to the eastern end of the equator to initiate the next event.

The theories of Suarez and Schopf (1988) and Battisti and Hirst (1989), which also have linear equatorial ocean dynamics at their core, provide a much clearer picture of how the ENSO cycle operates. More recent work along similar lines has expanded our understanding of this mechanism (Schopf and Suarez, 1990; Graham and White, 1988; Cane et al., 1990; Munnich et al., 1991; also see Cane and Zebiak, 1987), and a complete paradigm for the ENSO cycle can now be presented.

As in nature, let the main wind changes be in the central equatorial ocean while the SST changes are concentrated in the east. Then the surface wind amplitude, which depends on the east–west temperature gradient, varies with this eastern temperature. Further simplify by assuming that the eastern SSTs are principally controlled by thermocline depth variations. These variations are driven by the changes in the surface wind stress according to the linear shallow water equations on an equatorial beta plane. If the eastern SSTs are warm (thermocline high) then the wind anomaly will be westerly, forcing a Kelvin wave packet in the ocean to further depress the thermocline in the east thus enhancing this state.

However, this excess of warm water must be compensated somewhere by a region of colder water (shallower than normal thermocline). Equatorial dynamics dictates that this be in the form of equatorial Rossby wave packets, which must propagate westward from the wind forcing region. When they reach the western boundary they are reflected as "cold" equatorial Kelvin waves, which propagate eastward across the ocean to reduce the SST there.

Thus the original warm signal is invariably accompanied by a cold signal – but with a delay. This delayed oscillator mechanism accounts for the turnabout from warm to cold states.

To further appreciate the role of equatorial waves in sustaining the ENSO oscillation consider the state of affairs when the eastern thermocline and SST anomalies are near zero; for example, at the termination of a warm event. Then the wind anomaly must be near zero as well, so there is no direct driving to evolve the coupled system to its next phase. However, the previous warm event necessarily left a residue of cold Rossby waves in the western ocean, which eventually reflect at the west into a Kelvin wave which will reduce the SST in the east. The wind then becomes easterly and the cycle continues.

This paradigm may be distilled into a very simple system such as a single ordinary differential equation with a delay (Suarez and Schopf, 1988; Battisti and Hirst, 1989) or a recurrence relation in a single variable (Munnich et al., 1991). Perhaps the simplest version is that of Battisti and Hirst (1989):

$$\frac{\partial T}{\partial t} = -bT(t-\tau) + cT(t)\,. \tag{18.1}$$

They derived this equation as well as values for the parameters b, c, and τ, from Battisti's (1988) version of the ZC numerical model (Zebiak and Cane, 1987). Here T is the SST anomaly in the eastern equatorial Pacific, and c is the sum of all the processes that induce local changes in T, including horizontal advection, thermal damping, anomalous upwelling and changes in the local subsurface thermal structure (including local wave effects). The b term accounts for the effect of Kelvin waves generated at the western boundary as the reflection of Rossby waves; τ is the delay associated with this reflection process. Growing, oscillating solutions to (18.1) – ENSO modes – exist when $b\tau > \exp(c\tau - 1)$, a relation which holds for the parameters characteristic of the numerical model (see the Appendix to Battisti and Hirst, 1989).

Though other mechanisms can give rise to unstable oscillations in coupled tropical models (e.g., Hirst, 1986; 1988; Neelin, 1991; and see Sec. 18.6), it is generally accepted that this paradigm accounts for the behavior of the numerical models discussed above, as well as that in the higher-resolution coupled GCM which exhibits an ENSO-like oscillation (see Sec. 18.6). It is more difficult to establish conclusively that it operates in nature. It is consistent with the refill idea described above, which is supported by data (Wyrtki, 1985, and the additional time series available in the *Climate Diagnostics Bulletin* of NOAA). The role for western boundary reflection is further supported by the semi-empirical studies of Zebiak (1989) and Graham and White (1990). Finally, the ZC coupled model, in which this mechanism is clearly operative, has demonstrated the ability to predict warm events a year or more in advance (Sec. 18.5).

While the restriction of these models to the tropical Pacific region serves to bolster Bjerknes' emphasis on this region, it also renders them incapable of simulating the global consequences of ENSO. What is perhaps more troubling is the inability of the paradigm to account for the changes in the western equatorial Pacific preceding the warming in the east (see the discussion in Cane et al., 1990). More generally, the SO is observed to exhibit behavior distinct from El Niño (viz Fig. 18.1), and this too is not reproduced. These tropical Pacific omissions suggest that connections important to the ENSO cycle may have been overlooked.

The observed ENSO cycle is not regular, and some of the models share this feature (e.g., Figs. 18.3, 18.11). Nonetheless, the cause of the observed aperiodicity remains an unsettled issue. The results from Battisti's (1988) model and the experiments of Schopf and Suarez (1988) suggest that it is solely due to noise; that is, atmospheric or oceanic fluctuations distinct from the ENSO cycle. On the other hand, the low-order ENSO model of Münnich et al. (1991) produces aperiodicity, doing so rather readily if a seasonal modulation is included.

Experiments and analysis with ENSO models have demonstrated very strong sensitivities to rather small changes in parameter values. In the anomaly models some of these changes are equivalent to changes in the mean background state. Since a greenhouse warming will alter this state, the implication of such sensitivity is that the characteristics of ENSO will be changed. There have been a few experiments to explore this possibility (Zebiak and Cane, 1991), but inferences must be highly tentative in deference to our limited confidence in the ENSO models and to the great uncertainties as to the nature of greenhouse induced changes. This area of research is likely to become quite active as climate modeling progresses.

18.4 Modeling of ENSO as a coupled system

Though other views were (and still are) available, by the early 1980s a basis for modeling ENSO had emerged which allows much of the daunting complexity of the full atmosphere–ocean system to be ignored. The Bjerknes–Wyrtki scenario is at the core of it, which means that though the consequences are global, the essential interactions between atmosphere and ocean take place in the equatorial Pacific. The crucial variations of SST result from ocean dynamics, not variations in heat exchange with the atmosphere. Furthermore, these dynamics are essentially linear and act remotely: equatorial Kelvin waves carry the message of a wind change in the central and western equatorial Pacific eastward to effect a change in SST in the eastern Pacific. The role of the surface heat exchange is to drive the circulation of the tropical atmosphere, including the surface wind stress so crucial to the coupling. This atmospheric response can be largely captured by a steady-state linear model.