INTRODUCTION

In considering the theme of this congress, I have often thought about the evolution of marine science (oceanography) over the last thirty-five years. During these reflections, I marvel at the technical advances now available and sometimes wonder about our progress with substance. It seems that, as a community, we have pursued the science more or less as we wanted. Rightly, we wanted to understand marine systems in their pristine form before we tackled the impact of man on these systems. This is, however, no longer possible, as all major systems are under impact. Inevitably, an emerging role for science is that of assisting society in achieving sustainable use of natural resources. While this statement might not stir up much controversy among the present audience, the public would certainly question the implied urgency. Furthermore, exactly how science might ultimately realize this goal of assisting sustainability would surely be controversial among any group of scientists.

Human society is, out of necessity, reorganizing the vast network of interactions that connects it to the planet’s ecosphere. There are multiple solutions to this reorganization, of which the more human-friendly are a strong function of our own consciousness. Will science passively document the reorganization as it happens or will it proactively guide society to a favorable solution? The choice seems to me so important that I have chosen to insert a brief rationale for this changing role of science into this essay and to illustrate a specific class of interactions as an example of human influence on marine systems. For

SUMMARY: Human interference now represents an inextricable component of all major ecosystems. Whether this is through top-down overharvesting of ecosystem production or bottom-up alteration (deliberate or inadvertent) of the abiotic conditions, the planet’s ecosphere is in a vicious degradation cycle. For our economy to shift from exploiting to sustaining the natural systems, the solution, if there is to be one, will involve incorporation of the value of natural capital into the economic and political feedback loop. For the science sector, this will involve developing methodologies to evaluate the non-linear and behavioral dynamics of entire systems in ways that can be coupled with economic models. One essential characteristic of systems science involves the interactions between internal components and external systems. Thermohaline circulations and their feedback loops illustrate a class of such interactive pathways. Examples from the Arctic, Mediterranean, and the US East Coast along with some of their associated ecological impacts are reviewed. Understanding how thermohaline interactions provide stability to the marine biotic environment and under what conditions this stability could be destabilized is a fundamental step toward evaluating the non-linear response of marine systems to anthropogenic stress.

Key words: thermohaline, circulation feedback loops, natural capital, biocomplexity, Arctic, Mediterranean, Pamlico Sound.
this I will review how and why thermohaline dynamics maintain environmental stability and provide major interactive pathways between the abiotic and the biotic components of marine systems. Finally, in the last section I will try and coalesce these seemingly disparate topics.

A CHANGING ROLE FOR NATURAL SCIENCES

For those of us in touch with scientific literature, there is ample evidence concerning the degradation of natural systems. From available data, it is relatively easy to understand that the global rate of degradation exceeds that of regeneration. Further, understanding that regeneration is itself a function of the degradation, leads one to appreciate the condition of an exponentially declining sustainability. Recognizing that the force behind degradation is due to an unreasonable and uncontrolled economic exploitation, leads one to the conclusion that our market economy and growing population will cause a fatal collapse in the natural support systems on which our society depends. Finally, the logic of this sequence leads to the conclusion that the environment should be an integral part of the economy—not an externality to it.

A solution that I find tenable involves extending the role of science to include natural capital assessment. I believe there is evidence that this evolution is already occurring, but in a much too weak and fragmented manner. It needs a discussion within the science community and with the public. This extension will require scientific contribution beyond the so-called applied science of “impact statements” and “status and trends analyses” and into the arena of the simulation of complex, nonlinear, behavioral systems. It will also involve a change in how science interacts with society.

Natural Capital

Adam Smith based his economic theory on the availability of three ingredients: Labor (manpower), Capital (money) and Land (resources). His thesis was that the profit motive would drive a self-regulating market and increase the overall wealth of the community. Unfortunately, he assumed that the resources were unlimited and therefore gratis. The fact that 200 years ago natural resources did appear to be infinite may excuse Smith’s lapse. Now, however, there is no such excuse, and resources are still omitted from the supply-and-demand feedback loop where price optimizes the interplay only between Labor and Capital. In fact, the primary objective of the modern economy has been to convert natural resources into financial resources, using the argument that the process was reversible under a “substitutability” concept. In the 1970s, Daly (cf. Daly, 1991) proposed that the economy should be re-structured to a ‘Sustainable Economy’ by including an accounting of resources (Natural Capital) which, when incorporated into the supply and demand loop, would optimize the use of resources and minimize their degradation and depletion (Fig. 1). Today, having reached the global limits to our resource base, the assumption of infinite resources has become a fatal flaw.

Natural Capital implies both the principal, which is the use derived from ecosystems such as the pollination of crops or the filtration of freshwater, and the interest on the principal, which is harvesting of their yield such as the trees from forests and the fish from the sea. In a century of scientific observations, we have come to realize that natural systems are neither necessarily reversible nor inexhaustible. Yet Huxley’s quip, that the fish in the sea are inexhaustible, still rings in the ears of governments looking to subsidize their country’s participation in the tragedy of the marine commons.

Because the degradation of natural systems is considered an economic externality, the degradation of their function and yield exerts no force in the market economy. Without credible quantification and insertion of their values into the economic equation, any attempts to conserve the natural systems will be managed through legislative regulation and be necessarily confrontational with economic priorities. That is, there are two main alternatives: let the economy do it or let the governments do it. The former offers a much stronger self-regulatory dynamic (cf. Henderson, 1999; Hawkin et al. 1999).

Recent assessments of the Natural Capital value of the world’s ecosystems (cf. Costanza et al. 1997) and their rates of degradation (cf. Vitousek et al. 1997) has underlined the urgent need to better quantify the goods and services derived from Natural Capital in economic terms. Costanza’s evaluation of global renewable resources arrived at a figure of 30 to 55 trillion dollars, which was well above the Global National Product of $18 trillion in 1997. The message is not so much about the absolute value, which in many cases like the atmosphere, is ines-
timable. The message is how much will it cost us to live without, or to find alternatives for, these goods and services as they diminish. Vitousek’s article helps quantify the extent of human domination of natural systems and their rate of degradation. And Vitousek estimates that man’s sequestration of terrestrial primary production is \(\sim 45\%\).

When viewed from the point of view of global resource wealth, the per capita decrease is alarming. With resources declining annually at \(\sim 2\%\) and population increasing at 1.4\%, the per capita wealth is decreasing at 3.4\%. At these rates, we would have about two decades before global resource wealth will be half of its present value, a time scale far too short for a society which is having trouble understanding how its atmosphere—let alone the planet’s biodiversity—is changing, to confront and solve this problem. Meanwhile, the inertia of present policy is moving us toward further degradation.

**Role of science**

Once the goods and services of an ecosystem are known, even approximately, the relevant questions deal with how rapidly it is degrading, how resilient it is to further degradation, how it is linked to the degradation of other systems, and under what conditions it could restore itself. Estimates of these aspects can be translated into terms of risk assessment and probable costs that could steer economic use and policy in the direction of minimum damage. This is the social interface part of the new role.

The scientific part, which is perhaps more challenging, is learning to simulate living systems. Annual rates of degradation, as cited above, carry insufficient information to make such important predictions. We cannot assess and cost a function that we cannot understand. We are witnessing an increasing incidence of "environmental surprise"—systems collapsing or changing without formal scientific warning (cf. Bright, 2000; Myers, 1995). These phenomena involve a discontinuity or abrupt shift in what was perceived to be a stable system caused by the supersession of one internal dynamic over another, as for example the sudden loss of the Jamaican coral reefs (Hughes, 1994) or the sudden shift in bottom water production in the Eastern Mediterranean (below). They also occur when the impacts of multiple forcing functions combine synergistically to cause damage much greater than anticipated, as for example, the change in the phytoplankton community in the Danube plume due to increases in nitrogen and decreases in silicate (Humborg et al. 1997) or the record flooding of Hurricane Floyd in North Carolina due to climate variability and land-use practices (below). Addressing these problems
will require expansion of our methodologies to include an updated version of systems science, the new concepts of biocomplexity, and better observational means of monitoring the interactions between systems and their components.

“Biocomplexity” refers to a set of theories that modify, supplement, or limit our understanding of how living systems originate, organize themselves, and interact with the abiotic environment. According to this perspective, living systems can be considered deterministic and reversible only under limited conditions, otherwise they are indeterminate and irreversible. Much of this is explained by the fact that the abiotic environment is never in steady state and thus the living system is likewise in a ‘continuum of organization’ as it adapts to the variability in the mass, energy, and information fluxes that nourish and control it. The reorganizations of living systems do have constraints, (Lotka’s Maximum Power Law, for example (cf. Odum, 1983)), and some aspects like pattern are approximately repeatable, but we know little of the uncertainties of thresholds governing the bifurcation points that result in a change in organizational state (cf. Rapport and Whitford, 1999).

THERMOHALINE INTERACTIONS

Thermohaline interactions represent a very important abiotic connection for marine ecosystems. On a basin scale, they buffer the marine environment against atmospheric variability, and on a sub-basin scale, they act to stabilize the environment through recirculations that are beneficial to biological systems. For the reader unfamiliar with thermohaline circulations, I will summarize the principles needed to understand how thermohaline motions promote important ecological interactions and how they are self-regulating.

Thermohaline interactions involve inputs and exchanges that affect the buoyancy (density) of the surface water. Solar radiation results in a direct input of heat to the surface layer. Heat and water vapor are directly exchanged with the sea surface. Land runoff of freshwater is an indirect input from the atmosphere, but also dependent on the hydrogeological characteristics and the land use in the catchment basin. When the land receives the water evaporated from the sea, it should be considered as a partial exchange. Momentum is also an important atmospheric exchange that affects buoyancy indirectly through the displacement of surface waters. The formation and melting of sea ice represent a lagged exchange of heat and water vapor.

The salt in marine systems differentiates them from freshwater systems in their thermohaline circulations, biogeochemical cycling, and biological habitats. In all of these aspects, salt can be an important control variable. Here we mention it only with regard to its atmospheric interaction. While the sea-surface temperature has a strong interactive pathway with the atmosphere through its control of latent and sensible heat fluxes, the surface salinity has no direct interaction with the atmosphere. The atmosphere can change the salinity only through two independent exchange processes: precipitation and evaporation. The variable of salinity is not important to an atmospheric scientist, but is essential for an oceanographer. Consequently, salinity is important for coupled atmospheric-ocean models, particularly at high latitudes where the density is mostly controlled by the salinity, and numerical models must include the dynamics of freshwater and convective motions to achieve correct salinities (cf. Bryan, 1986). Convection is an irreversible process as far the atmosphere is concerned: that is, the parcel that lost water to the atmosphere sinks and is no longer available for dilution back to its original salinity. Convection, rain, and ice formation are all processes that can delay or displace the impact of strong atmospheric interactions and thereby contribute to the stability of thermohaline circulations by dispersing the buoyancy disturbance in space and time.

Thermohaline circulations

Thermohaline circulations are initiated when the buoyancy of surface water is changed at some location relative to another. This is a continuous process because the atmospheric variables responsible for buoyancy exchange vary at different time and space scales relative to the ocean such that there is only rarely a zero-buoyancy flux. Any local change in buoyancy affects the weight of the water column, which in turn affects the internal pressure field and induces circulation.

A brief example is that of an ice cube put into a tub of salt water. First, it bobs in the vertical as the ambient pressure field isostatically adjusts so that the ice floats high enough such that weight of the ice, supported above the water is, following Archimedes’ Principle, equal to the weight difference between the submerged ice and the salt water.
As the ice melts, a blob of freshwater will remain floating, slightly elevated, but no longer rigidly contained. The pressure field will induce small currents, which, along with molecular diffusion, will mix salt into the blob.

The pressure field is best understood if it is linearly separated into a barotropic component created by the dislevel of the sea surface and a baroclinic component created by the differential weight distribution under the surface. At the sea surface, the barotropic pressure gradient forces the melt water to move outwards from the most elevated point. Underneath, the baroclinic pressure gradient forces water to move toward the same point where the weight difference is a maximum. This compensatory process, provided by the available baroclinicity, leads to two important consequences: the vertical two-way flow that is characteristic of thermohaline circulations and the absorption with depth of barotropic disturbances that otherwise might have impacted the entire water column. In freshwater, the available baroclinicity would have been only a function of the temperature difference and hence these compensatory processes would have been much weaker.

If the sea surface is warmed or diluted, buoyancy is added, and a positive thermohaline circulation is initiated. If the source of buoyancy is far from boundaries, an anticyclone (northern hemisphere) would set up; if it is along the coast, a coastal boundary current would be formed; if it were inside a semi-enclosed basin, a surface outflow would occur. In all these cases, the deeper water will be forced to move underneath the surface water, while the surface water will be forced to move away from its source. The result of this estuarine shear is the process of entrainment that serves as such an important abiotic control for estuarine ecosystems.

If the sea surface is cooled or evaporated, buoyancy is extracted, and a negative thermohaline circulation is initiated. In this case, however, the affected water sinks to a subsurface level matching its density. If it reaches the bottom, it will move downslope on the right side of the basin and replace any resident deeper waters of lesser density. Typically, we find dense water-mass accumulations where there are bathymetric depressions and/or bottom-water currents along the western sides of the ocean basins.

Two other important differences in these circulations need to be mentioned. One involves the time-dependency of the buoyancy forcing. Addition of positive buoyancy by river runoff occurs in a continuous but variable form. On the other hand, the extraction of buoyancy, sufficient to cause strong convection, occurs seasonally with interannual variability. The second difference is that the surface water of the positive circulation entrains the deeper water that it moves over. This is a dissipative process because it weakens the pressure gradients that are driving the flow. On the other hand, the bottom-water flows are less dissipative because there tends to be less kinetic energy, less shear, and less density difference with the overlying waters. Thus, negative buoyancy waters are formed less frequently, but they also dissipate less and tend to accumulate more which gives longer residence times and the circulations that they drive more stability.

We can now summarize briefly the thermohaline processes that generate the nonlinear exchange for semi-enclosed regions (cf. Hopkins, 1999). I will use the term “basin” to mean the region and “strait” to mean its constricted connection with the external ocean. The following characteristics describe a positive circulation, with the appropriate words in parenthesis for the negative circulation case.

1. A basin is exposed to significant buoyancy addition (loss) due to differential rates of heating/freshening (cooling/evaporation), relative to that of the ocean.

2. The density of water columns inside the basin decreases (increases) making them have less (more) weight (PE) than a water column outside the strait. The resulting baroclinic pressure gradient at the depth of the sill forces water into (out of) the basin. This baroclinic force increases with depth, because the weight increases, and thus also does the resulting baroclinic flow. The length scale of the baroclinic pressure gradient is several times the length of the strait.

3. Contemporaneously, when the buoyancy is added (subtracted), the basin sea level rises (lowers) and the resulting barotropic pressure gradient generates an outflow (inflow) constant with depth.

4. This initial sea-level difference would be quickly resolved by the one-way barotropic outflow (inflow) were it not for the fact that, as the opposing baroclinic flow enters (exits), it raises (lowers) the sea level.

5. At the bottom, the inflow (outflow) is subject to frictional force that makes the flow decrease at the bottom. This force is proportional to the square of the velocity, the roughness of the bottom, and the bottom area of the strait. Thus the velocity profile has a maximum outflow (inflow) at the surface and a maximum inflow (outflow) above the bottom, as in
Figure 2. With greater friction the interface, between inflow and outflow, becomes shallower so that the two flows still balance.

6. The two-way exchange is constrained to conserve the volume of water in the basin such that the net exchange equals the internal water balance.

7. The buoyancy is not necessarily conserved, but over time it converges to a running mean condition.

8. The sea-level gradient can be independently influenced by tidal, wind forcing, or atmospheric pressure differences. This causes the sea level to oscillate about a zero-mean level. These generate an oscillating barotropic flow component that is superimposed on the thermohaline exchange. The associated water transport integrates to zero over time.

Three complicating situations affect the magnitude of the inflow and outflow. They are mentioned here because they enter in the discussion of the following examples.

9. The sea-level oscillations (#8) can increase the efficiency of the water-property exchange, by sloshing basin-water out of the strait (barotropic pumping) and not sloshing the same basin-water back through the strait.

10. The combined flow (thermohaline plus oscillating) can also be large enough such that the constriction in the strait reflects the flow, with the result that efficiency of the exchange is reduced (hydraulic control).

11. In wide straits, the vertical organization of the inflow and outflow gives way to a lateral organization. This allows for a greater exchange, which laterally recirculates inside the basin and can change the internal buoyancy through mixing.

Thus, the exchange solution is governed by a fairly simple non-linear relationship.

12. The thermohaline exchange is proportional to the force driving it and the force is reduced by the net exchange that it drives. The proportionality is determined by the characteristics of the strait and of the higher frequency oscillations driving the barotropic pumping.

Since the exchange depends on itself (via the force), it has an exponential dependency with time. Knowing the force, the bottom friction, and the buoyancy input, one can compute the net exchange. If the above complicating factors (#9, 10, and 11) exist, they must be accounted for and require additional observations to determine their effect.

Stability of the exchange

Item 12 functions to make both the volume and salt exchange self-regulating. The volume is actively conserved (#6) controlled by a simple sea-level feedback loop that is very similar for the two circulations. If we look at a perturbation in the sea level inside a basin, as for example that caused by an extreme rainfall event, the barotropic outflow increases in proportion to the change in sea level over a very short time scale. The excess water generates a one-way, barotropic supplement to the thermohaline exchange. As the excess volume decreases, so does the outflow until the exchange returns to its previous condition. Thus, we have an effective stabilizing feedback loop (FBL). The above oscillations (#8) are controlled in this manner. Of course, the basin’s barotropic response can be in resonance with an oscillating force and create amplifications (e.g. tides in the Bay of Fundy).

The buoyancy feedback loop is more complex, and we need to consider the two circulations separately. For the positive circulation, the most common situation is that of freshwater addition as found in coastal estuarine systems. The freshwater input, whether totally mixed into the internal waters or not,
creates an integrated density difference through the strait (or inlet) that drives a positive exchange. Therefore, a runoff event will increase the exchange with the ocean. The increased inflow of salty ocean water will gradually decrease the buoyancy gradient driving the exchange, and the system will return to its mean situation. This is, then, the stabilizing feedback loop, as sketched in Figure 3a. While it appears directly similar to the sea-level scenario described above, it differs for two reasons. First, the sea-level perturbation creates a one-way flow, and the buoyancy perturbation creates a two-way flow, which is slower in dampening the disturbance. Second, the time required for information about a change in sea level to reach the strait is much shorter than that required for information about a change in density. For example, the sea-level adjustment about a runoff peak will arrive quickly because it only involves vertical movement as the barotropic wave passes; however, the information about additional freshwater involves the transport of the freshwater itself to the neighborhood of the strait where the baroclinic pressure gradient is formed. Thus, the size and the configuration of the estuary tend to buffer the buoyancy signal and delay the time to re-stabilize the exchange.

For the negative circulations, the controlling feedback loop is illustrated in Figure 3b. If, during a
particularly cold dry winter, an excess of dense water is produced, it will more than replenish the negative buoyancy lost by the outflow during the previous year. The greater accumulation will increase the outflow and thereby cause more buoyancy to be imported from the external ocean. When the next cold season occurs, given the same atmospheric conditions, there will be relatively less dense-water production because there had been more opposing buoyancy input. In this way, atmospheric trends will be mapped into the amount of dense water stored in the system and control the compensating import of buoyancy. As the atmospheric conditions return to their mean condition, so will the exchange. There will be a lag and a sensitivity to this compensation that will depend on the reservoir of dense water stored in the system, the distance from the production site to the strait, and the buoyancy characteristics of the replacement flow (Table 1). Even without replacement, the stored deep waters can gain buoyancy and increase the probability of replacement through vertical diffusion processes and even geothermal heating; for example, renewal of the Jabuka Pit in the central Adriatic is virtually assured because the bottom water is invariably cooler and less saline than the overlying waters (Hopkins, et al. 1998a).

### Table 1. - The degree of buoyancy stability provided for by the replacement inflow. The abbreviations are: VS – very strong; S – strong; W – weak; VW – very weak

<table>
<thead>
<tr>
<th>INFLOW</th>
<th>Warming</th>
<th>Wetting</th>
<th>Cooling</th>
<th>Evaporating</th>
</tr>
</thead>
<tbody>
<tr>
<td>warm</td>
<td>VS</td>
<td>S</td>
<td>VS</td>
<td>W</td>
</tr>
<tr>
<td>cool</td>
<td>S</td>
<td>W</td>
<td>S</td>
<td>VS</td>
</tr>
<tr>
<td>fresh</td>
<td>VS</td>
<td>S</td>
<td>VS</td>
<td>S</td>
</tr>
<tr>
<td>salty</td>
<td>S</td>
<td>VW</td>
<td>S</td>
<td>VS</td>
</tr>
</tbody>
</table>

reverse them and completely change the thermohaline structure of a marine system causing a kind of thermohaline environmental surprise? Brief descriptions of three different systems will illustrate how thermohaline interactions process abiotic variability and how they function in controlling the marine environment.

### Pamlico Sound system

The Pamlico Sound provides an interesting example of a positive thermohaline circulation because it is a complex ecosystem involving a gradual transition from fresh-to-salt water habitats, through which there is also a transition in the type of thermohaline control. In addition, the Pamlico Sound is being increasingly impacted by anthropogenic stresses, for example, through inappropriate land use (flooding), inappropriate agricultural practices (nutrients and pesticides), groundwater overdrafting, habitat destruction (wetland development), beach erosion on the barrier islands, and over-exploitation of fisheries resources. To appreciate the natural capital of the system and its vulnerability, it is enough to say that it the Pamlico System is the second largest US estuary, a major nursery and fishing area for the Atlantic coast, one of the fastest developing US coastal areas, and one of the most frequently affected by tropical storms.

### System

The Pamlico Sound and the adjoining Albemarle Sound constitute a shallow (< 7 m) body of water trapped behind the barrier island chain of North Carolina (Fig. 4). Because of their shallow depth and nearly closed geomorphology, the sound system is characterized by low flushing, small tides, and strong wind mixing (e.g. Pietrafesa and Janowitz, 1988)). The four major rivers that provide the land
drainage to the system discharge through four partially stratified estuaries. These estuaries act to process much of the nutrient and organic matter before it reaches the sounds. Thus, the sound system acts as a holding basin for the freshwater and for further biogeochemical processing of discharged organic matter and nutrients. The sounds serve as sites for deposition and re-working of sediments. Significantly, they provide large, brackish-water habitat for all or portions of the life cycles of many commercially valuable species. By regulating the flushing times, determining salinity, and providing a conduit for the recruitment and emigration of many estuarine organisms, the thermohaline exchange with the Atlantic Ocean controls the environment by buffering the oceanic and riverine influences. The flushing time provided by the ocean inflow is ~3 mo whereas that by the runoff is ~12 mo.

The Pamlico is connected to the Atlantic Ocean by three inlets (Oregon, Hatteras, and Ocracoke), to the Albemarle Sound northward through the wide Roanoke Channel, and to the Core Sound southward through a much smaller channel. Its two watersheds are similar. The Neuse River is a little larger and has a greater urban area than the Tar River, which drains a relatively more agricultural area. The combined runoff is ~9 km³/yr. Both these rivers have narrow, partially stratified estuaries that are similar in size and degree of eutrophication (e.g. Paerl et al., 1998; Stanley, 1997). Both have frequent summertime anoxic events with frequent fish kills as well as occasional outbreaks of harmful algal blooms (e.g. Glasgow and Burkholder, 2000). The adjoining Albemarle Sound is also fed by two major rivers, the Roanoke and the Chowatan, which deliver another ~13 km³ of freshwater into the Pamlico near the Ore-
gon Inlet. The Neuse and Tar Estuaries have been well studied, but the Pamlico Sound and the Albemarle Sound, with its river systems, have received much less attention.

**Thermohaline exchange**

We have constructed a preliminary thermohaline exchange model as part of an attempt to study the entire system (Hopkins and Molina, submitted). The Pamlico and Albemarle Sounds were each considered as volumes of homogeneously mixed waters. They receive the freshwater from their respective rivers, the signal of which is delayed and spread out (over 1-2 days) as it is incorporated into each of the Sound’s volumes. The exchange with the ocean is represented with a single inlet having a cross-sectional area equal to that of the three actual inlets.

A non-linear process modeling software (EXTEND™) was used to simulate the exchange at a daily time-step. The only period where both input data and in-situ calibration data were available was from January 1998 to September 2000. The runoff data came from the US Geodetic Survey, the atmospheric data from the US Weather Service, the coastal salinity of the Atlantic from the Duck Pier Facility, and the calibration salinity-data from the western part of the Pamlico Sound from Joe Ramus, Duke Marine Lab. The model keeps track of the density, the sea level, and the freshwater in each of the sounds and calculates the respective exchanges according to the sequence outlined above. The two unknowns are the interfacial mixing (function of Richardson’s Number) salinity in the inlet and the friction (function of square of bottom speed) in the bottom layer. Because they both vary as a function of systems variables and have dissimilar effects on the solution, they can be calibrated quasi-independently. The ocean salinity south of Cape Hatteras has a different variability and amplitude (by ~ 0-3 ppt) than that to the north. In addition, the inflow is greater at the southern inlets during southwesterly winds, common in summer (Xie and Pietrafesa, 1999). This uncertainty was partially simulated by adding a seasonally varying correction to the amplitude of the Duck-Pier salinities (Hopkins and Molina, submitted).

**Stability**

In the Pamlico Sound, the thermohaline exchange acts to stabilize its brackish water habitats by buffering the internal salinity relative to the variability of the runoff and ocean salinity. There are several basic questions concerning the capacity of this buffer during extreme storm events. For example, the resilience of the ecosystem may be exceeded under the impact of introducing large quantities of water, terrestrial matter, and nutrients into the system. While the physical system may be fairly robust, its interactions with the biological system may not be. Ecological recovery is particularly in question under the present trend for increased incidence of hurricanes, as discussed by Paerl et al. (2001). Understanding how and when these interactions switch from stabilizing to destabilizing for any given system is a fundamental step.
toward evaluating their non-linear response to anthropogenic stress.

Figure 5 shows the salinity response of the Pamlico Sound during a period which included the record flood of Hurricane Floyd on 15-16 September 1999. The storm waters completely overwhelmed the system and continued to freshen the system throughout the nine-week period of elevated land drainage. From the minimum salinities in mid December, it took until the following June to re-establish normal salinities. The hysteresis between the response and the recovery occurred because the freshening was forced by the event, whereas the recovery was a result of the stabilizing feedback loop. For shorter events, the recovery time can be quicker than the duration of the event, as seen with the discrete runoff peak at day 240 in Figure 5. This happens because the restorative action is strongest during the event and tapers off exponentially as the buoyancy gradient force decreases during recovery. Very important is that the recovery time of the salt content does not imply an equivalent flushing time for the particulate organics, nutrients, or flocculated toxins all of which react within the pelagic system and portions of which become stored in the benthic system (ibid.).

The Pamlico System offers an example of how the thermohaline control could make a sudden shift to a higher mean salinity as a consequence of a substantial breach in the barrier island system. The barrier islands are vulnerable due to beach development, increased storm frequency, and sea-level rise. A large breach could change the wave environment and hence accelerate wetland shoreline erosion, and any breach would raise the salinity, possibly beyond tolerance, for many of the brackish water organisms. This is an example of a geological (and biological) thermohaline interaction that would greatly alter the natural capital of the system.

Estuarine interaction

The same thermohaline dynamics control the circulation in all segments of the estuarine system, but they are differently manifested in the case of vertical stratification, and the interaction with the biological system can be quite different. As an example, the Neuse Estuary is a partially stratified segment that is embedded in a larger well-mixed system (Figs. 4 and 6). The upper reaches are river-dominated, slightly brackish, vertically mixed, and input to the upper layer of the Neuse Estuary. At the head of estuary, the cross section widens and deepens such that there is not enough kinetic energy to maintain mixed-water columns. At the mouth where the estuary connects with the Pamlico Sound, the cross-section widens even more, and the situation reverts to a wind-mixed regime. The geomorphology is a strong control, but not exclusive. Consequently, the landward and seaward boundaries can extend in either direction depending on the amount of kinetic energy available (runoff or wind) and the buoyancy differential with adjacent segments.

The thermohaline circulation of this segment, like any other, has the same stabilizing feature regarding its exchange: that is, an increased runoff induces an increased inflow. However, the fact that it is connected in series with two other estuarine systems limits its stabilizing capacity. For example, with a large runoff event, the input from the river-segment freshens, the outflow to the Sound freshens, the inflow from the sound freshens, and the entrainment of salt...
decreases, making the outflow to the Sound even fresher, etc., until the upper boundary collapses and the Neuse Estuary loses its partially stratified condition. Note that during such an event, the exchange between the Sound and the Ocean brings in more salt but, as was seen in the above example, the salt recovery time for the Sound is long.

Biotic connection

Aspects of abiotic-biotic interaction can be illustrated with several examples. Because the runoff discharges both freshwater and nutrients to the system, the freshwater variable becomes a catalyst for abiotic-biotic interaction. An increase in freshwater input increases primary production and also it increases the thermohaline circulation. As the phytoplankton biomass increases, it advects down the estuary; as it dies, it sinks into the lower layer; as it returns up the estuary, some of it regenerates. The regenerated nutrients are then entrained into the surface layer and help sustain the surface-layer production. Were it not for this thermohaline interaction, a pulse in the river flow would generate a bloom that would be advected out of the geographic location, as occurs in a river. In the opposite case of a decrease in the freshwater input, the loop process slows and shortens. Thus, the thermohaline circulation helps stabilize the geographic location of the primary production and reinforce its growth.

The bottom oxygen is also indirectly stabilized by the same loop. In the above example of a runoff event, the increased production and subsequent regeneration requires an increased demand for oxygen in the bottom layer, which is then compensated for by an increased advective input of oxygen from the mouth of the estuary. Here we have an example of how the stabilizing effect is restricted over a narrow range of events: for example, a larger runoff event may stimulate a greater production that would extend past the boundaries of the estuary and eventually reduce the oxygen levels of the bottom inflow.

A third, related, example concerns the bottom organic loading of particulate matter. At some point near the head of the estuary, there is a flow convergence where the landward, bottom-layer flow of the estuary equals the seaward flow of the riverine mixed segment (Fig. 6). This is the point of furthest consistent advective penetration of salty water and is where the ‘salt wedge’ forms as the isohalines of the partially stratified estuary intersect the bottom. This bottom flow convergence generates a “turbidity maximum” where, due to the null-speed at the bottom, particulates can settle out faster than they are resuspended. Note, the location of the null-point migrates, as does the upstream boundary of the estuary. The effect is to accumulate organic-rich sediments on the bottom. For a certain middle range of the thermohaline and primary production, this carbon accumulation can represent a stabilizing storage of labile carbon and a site of nutrient generation. For slower or faster dynamics, it can become significantly destabilizing for the ecosystem. In the extreme runoff event sequence (above), much of the material is resuspended by the kinetic energy of the runoff and transported out to the Sound where, in addition to the increased organic load of the river, it increases the POC, turbidity, BOD, etc. During a summer period of low runoff, higher temperatures, and less advective replacement of oxygen, this accumulated organic matter drives down the bottom-layer oxygen and creates the destabilizing condition of hypoxia for the system. It also provides a site for denitrification, promoting a loss of nitrogen from the system.

In the sequence of the above three examples, the primary interactions were among the freshwater, nutrient supply, and primary production. The connections with oxygen and sediments expanded from this first interaction. This partially demonstrates how internal interactions propagate within the system and take on other interactions. In the end, the signal of the abiotic event is either buffered or enhanced as it moves through the network of connections. From the point of view of the physical thermohaline stability, the system is able to recover when exposed to extreme events. Apart from the physical resilience, one might question the resilience of certain organisms exposed to rapid salinity changes or long periods of recovery. Recent findings have shown the flood-related loading to the system of terrestrial organic matter caused an impact of longer duration (Paerl, 2001). Critically important to the natural-capital assessment of the system would be the threshold switches that cause abrupt changes in the network of interactions that stimulate internal reorganization.

Mediterranean system

The Mediterranean is the reference system for negative thermohaline circulations (cf. Hopkins 1978). The Mediterranean marine, atmospheric and terrestrial environment defines a particular culture that has deep roots in human history. The threat of
losing this environment haunts the cultural consciousness. It is a mute question whether this derives from the dramatic paleo-environmental changes, by analogy with the neighboring Black Sea, or from the modern trends of climate change, sea-level rise, loss of fishery, and loss of a natural aesthetic. Hence, there is a strong use-value associated with preserving the Mediterranean’s marine environment. Fisheries and loss of the coastal zone assets are forced mostly by socio-economic forces. Sea-level rise and changes in thermohaline circulation that could result in hypoxia are forced by climate trends. An often asked question is whether the Mediterranean could serve as a scaled-down model for the study of climate change. Another question regards whether or not the Gibraltar exchange is a sensitive indicator of climatic trends (e.g. Garrett et al., 1990; Hopkins, 1999.) Recently, large changes observed in the Eastern Mediterranean (EMED) concerning the origin and distribution of the intermediate and deep waters (Malanotte-Rizzoli et al. 1999) have raised questions about interactions between the thermohaline circulations of the Mediterranean’s sub-basins.

System

The Mediterranean’s thermohaline circulation can be characterized as having a primary thermohaline cell circulating west to east in the upper layer above the depth of the two major sills (<400) as well as less-active cells in the deep basins below. It also has a secondary south-to-north circulation that combines with the principal west-to-east circulation. This feature derives from the basic zonal orientation of its two major cells and where each of them extends sufficiently in the meridional direction to experience significant gradients in atmospheric forcing. This situation adds complexity to its thermohaline circulation and sensitivity to atmospheric variability.

The Mediterranean’s general thermohaline stability is marked by its predominantly haline control of dense-water production, by its large capacity for dense-water accumulation, by its different production sites, and by its limited exchange with the Atlantic Ocean. The net loss of ~1 m/yr of freshwater dominates its capacity for dense-water production. In fact, 75% of the density increase is due to salinity increases, with cooling causing the remaining 25% (Hopkins, 1999). The thermal fraction plays an important role in promoting deep convection, with the result that most of the bottom-water products (WMED and Adriatic) are less saline and cooler than the overlying waters, which gives them greater diffusive instability.

The Mediterranean case well illustrates that the dense-water production process can be continuous, but only expressed during the cold season, and that the process is usually spatially extensive, but only completed at specific sites. During the summer, high evaporation rates increase the salinity of the surface waters, particularly in southern portions where there is less runoff and frequently drier air along the North African Coast. This haline preconditioning increases the surface salinity and facilitates subsequent dense-water production through heat extraction. The cooling process can be divided into two phenomena, ubiquitous and site-specific.

The autumnal cooling is a ubiquitous process. If it weren’t for the meridional variation, the process would be much more reversible. In the north, the loss of summer heat occurs sooner and continues longer than to the south. Any convection that overturns waters deeper than the summer thermocline produces the first dense-water product of the season, which is sometimes referred to regionally as Local Winter Water (LWW). The LWW is produced practically everywhere within a cooling length scale of cold northerly winds. It does not generally accumulate interannually. LWW is subject to mixing with adjacent layers and is not included in accumulation volume. However, the LWW is a very important part of the thermohaline response, and, in fact, it initiates the seasonal thermohaline signal. Because of the differential phasing and intensity of its production between adjoining basins, the baroclinic pressure gradients through straits change accordingly and stimulate internal exchanges, for example, that between the Tyrrenian and Ligurian Seas (cf. Hopkins, 1990a; Astraldi, 1990).

The intermediate water products, which reside below the LWW, drive the primary thermohaline cell for the Mediterranean. The most important of these are those generated in the EMED, the Levantine Intermediate Water (LIW) and, recently, the Cretan Intermediate Water (CIW). They are generally identified by a vertical temperature and salinity maximum. The intermediate water masses do not accumulate in the sense of being bathymetrically trapped, but they can persist for more than one season depending on their advective travel time through the basin. The passage can take six months to a year (La Violette and Manzella, 1990) and exposes the LIW to mixing with the overlying waters and

THERMOHALINE FEEDBACK LOOPS 243
entainment with the underlying waters. As it moves westward, the continual entainment of underlying bottom-water makes its water-type converge towards the deep-water type resident in the Alboran (Wust, 1961). Because the exchange through Gibraltar is determined only by the Atlantic-Alboran baroclinic pressure gradient, the exchange is responsive only to the internal circulations that deliver the LIW to the Alboran and not to the production in the Levantine (Hopkins, 1999).

As winter progresses, the greatest production of dense water occurs at locations where the circulation dynamics bring preconditioned waters to the surface and where the surface receives maximum exposure to cold, dry winds: the Gulf of Lyon producing bottom water for the WMED, the Adriatic and Aegean producing bottom water masses for the EMED and the northern Levantine producing the LIW. The bottom-water products sink beneath the level of the sills to the bottom and contribute directly to the deep-water accumulations. As it sinks, the new water spreads along isopycnals ultimately towards the basin sill, giving western and southern components to their movement. [The Tyrrenian is a notable exception with bottom waters entering from the southwest, mixing with the LIW and exiting above the bottom water as a less-dense water mass (Hopkins, 1988)]. Thus, accumulation basins exhibit slight vertical gradients in density and each year’s crop of new production displaces the less dense, overlying waters. This displacement and other upwelling processes link the deep-cell circulation with that of the upper-layer circulation above them and allow their exit from the basin. That is, the bottom-water products have a much longer residence time than those above, and practically any annual signal in their production is obscured in the upper thermohaline cell. The two sub-basins of the Adriatic and Aegean have their own accumulation basins, such that their products accumulate before exiting over their respective sills before reaching the EMED.

Of interest is the double thermohaline circulation of the Northern Adriatic Sea (Hopkins, 1999b), which is noted for providing for the Mediterranean with the coldest, least saline dense-water product and the largest source of riverine discharge. The fact that the discharge is large and the bathymetry of the Italian coast is smooth facilitates the capture and export of ~85% of the freshwater in a geostrophically trapped boundary current, which effectively prevents the freshwater from interfering with the positive circulation. Variations in the amount of freshwater mixed into the interior act as a control on the dense-water production and make the product less saline than its source water. This variability is mostly controlled by the incidence of the Sirocco winds from the southeast (Sirocco) and affects the stability of the negative portion of the thermohaline circulation. The dense-water production is also controlled by the incidence of the cold, dry winter Bora winds from the northeast (Orlic, et al. 1994). The dense water formed flows southeast along the western side and into the deep basin of the Southern Adriatic where it combines with locally produced dense waters before exiting over the Otranto Sill to the EMED deep water.

Feedback loops and exchange

The controlling feedback loop for the Mediterranean is direct (Fig. 3b): the more buoyancy is extracted by the atmosphere, the more buoyancy is imported from the Atlantic. The stability of this loop is assisted by the greater buoyancy of the Atlantic replacement water from a both heat and salt point of view (Table 1). The stability of the exchange is greatly increased by the large accumulation of dense water. Consequently, although its response to atmospheric variability is direct, the final expression at Gibraltar is not very responsive to internal production of dense water. The response is also complicated by the length of time required for the signal of different water-mass productions to reach the strait. For example, this gives Sicily roughly a half-year lag and Gibraltar a one-year lag with respect to LIW production and offers some possibility for resonant coupling with seasonal or annual variability in the atmospheric forcing.

If the water columns above the sill are represented as a two-layer structure consisting of the mean North Atlantic and Mediterranean water types, the interface would be at ~120 m or equivalent to an accumulation level of 180 m above the Gibraltar sill. The half-life for this accumulation to drain is ~9 yrs which therefore makes the outflow insensitive to annual variations in buoyancy extraction (Hopkins, 1999). The variability observed in the Strait is related to independently generated transitory motions (wind, tides, and atmospheric pressure differences) and the auto-regulating effect of the thermohaline-feedback loop. Despite these transients, the net exchange is determined by the height of accumulation in the basin and converges to its mean condition. For the range of experimental results and with
little variability in the Atlantic water type, the force and the exchange have a fairly linear relationship as shown in Figure 7.

Another higher frequency process of auto-regulation occurs in the Gibraltar Strait. Through the mechanism of barotropic pumping, the independently forced sea-level oscillations play an important role in the outflow (~33%) by sloshing the Mediterranean water over the sill (Hopkins, 1999). This sloshing effectively evacuates the Mediterranean water in the western Alboran and, therefore, the thermohaline exchange slows down until the water is supplied from the west. Without this mechanism to make the water-mass outflow more efficient, the internal level of accumulation would be significantly higher.

The exchange through the Strait of Sicily behaves similarly, but with some complications. The passage is wider and therefore can support surface recirculations. The passage is also longer and contains an accumulation basin in the Pantelleria Trough for the westbound dense waters. The strait is too long for most sea-level oscillations to generate a barotropic pumping similar to that at Gibraltar. The exception may be some sloshing of denser waters into the holding basin (Ribera, pers. comm.). The pressure gradient through the Strait that controls the LIW outflow is between the Ionian and Tyrrenhian, which is, in turn, controlled by the gradient through the Sardinian Channel, etc. Again, the consistency of Sicilian outflow is due to the large accumulation in the EMED. However, variability in the outflow can be due to the seasonal phasing of internal pressure differences between adjoining waters in the Tyrrenhian and Ionian, due to the variability in supply of LIW caused by the circulation in the Ionian and due to circulations internal to the Strait.

Long-term stability

Consequently, the Mediterranean Sea has a very stable thermohaline condition. Even under an atmospheric warming scenario in which the weather patterns no longer generated the continental, cold-air outbreaks responsible for the bottom water production, the upper thermohaline cell would still continue to circulate as long as the negative buoyancy balance persisted across its straits. It is unlikely that the Northern Atlantic would become saltier than the Mediterranean. Hence, the intermediate-water production would continue, and the upward diffusion of negative buoyancy stored in the deep waters would contribute to forcing the exchange for a long time. A warming trend would increase the evaporation and result in a warmer, saltier LIW product. A concurrent trend in increased precipitation would tend to offset this increased evaporation except for areas like the Aegean and Levantine basins where the freshwater runoff is decreasing.
The later is a result of the increased offstream water use of the Nile and Black Sea rivers over the past decades. The net consequence of these present trends is indicated with the observed warmer, saltier bottom water in the WMED, as reported by Bethoux et al. 1990. Only a major change in precipitation patterns, particularly over North African and the Middle East, coupled with warmer winters, might overcome the buoyancy budget and convert the Mediterranean to a positive thermohaline circulation like the Black Sea. A shutdown of the deep ventilation would have great biochemical consequences; for example, a much-reduced renewal of deep oxygen and a much-reduced upwelling of nutrients to the upper layer.

Internal interactions

Although the stability of the primary Mediterranean thermohaline circulation is robust, its internal interactions may not be. An excellent example of ‘environmental surprise’ has already been observed in the EMED. Historical oceanographic evidence has always cited the Adriatic Dense Water product as the source of the bottom water volume in the EMED. It was clearly identifiable by its lower salinity and colder water type (Pollak, M.J., 1951; Hopkins, 1978), which first stored itself in the deep Ionian and then in the eastern Levantine Basin. During the late 1980s, this role was assumed by the waters of the Cretan Basin which had increased in density, as has been clearly demonstrated by Roether, et al. 1996; Malanotte-Rizzoli, 1999). Previously, the Cretan waters contributed only in small volumes through the straits of the Cretan Island Arc (Miller, 1974) because there is only a slight density difference with the adjoining waters in the Levantine Basin. Their increase in density, together with the reduction in the amount of the Adriatic Dense Water product, made the difference.

Fig. 8 – The double thermohaline circulation of the Northern Adriatic Sea. The positive thermohaline, coastal circulation is driven by the runoff and assisted by the Bora winds and weakened by the Sirocco winds. The negative thermohaline, interior circulation is driven by winter cooling and is also assisted by the Bora winds and weakened by mixtures of freshwater caused by the Sirocco winds. The negative circulation can have a positive feedback loop by which weak interference on the part of the freshwater can increase the dense-water production, and vice versa. The overall control of the negative circulation is that of Fig. 3b in which in which the replacement flow and the production dampen variations, as in Fig.3b. However, this loop is only moderately stable because the replacement flow is warmer and only slightly more haline than the dense water product (Table 1), and its negative buoyancy potential can be weakened by mixtures with the freshwater of the coastal outflow (question marks in diagram).
The causal relation has not yet been documented. However, it is clear that the atmospheric warming trend and the reduced freshwater runoff are implicated. The increase in the Cretan Deep Water density derives from a salinity increase. Two obvious causes are the cessation of the Nile discharge since 1965, which had reduced the salinity of the input to the Aegean Sea from the southeast (Lacombe et al. 1958), and the reduced freshwater inflow from the Black Sea (Kempe et al. 1990). Likewise, atmospheric trends in the Northern Adriatic have resulted in warmer winters with fewer incidents of the cold Bora winds that are essential to local dense water formation (Gacic, et al. 1998). At the same time, a trend has developed toward more frequent, large, runoff events in the fall season (Russo, et. al. 2001) that have acted to inhibit dense water production (Hopkins, 1998).

The interactions created by the double circulation of the Northern Adriatic (Fig. 8) have implications for both the physical and biological environments. The replacement waters for both circulations are of the same origin deriving from the surface and intermediate waters to the southeast and ultimately from the Ionian. Thus, the increased salinity of the new CIW is impacting the source water for the Adriatic dense water product (Klein, et al. 2000), providing an example of cascading thermohaline interaction. The negative circulation requires a seasonally decreasing replacement volume as the dense water drains out, which can last through summer. The positive circulation requires a variable replacement dependent on the runoff. Its volume is several times that required by the dense-water outflow. While this combined flushing is vigorous and acts to dampen the biological response to the strong nutrient loading of the system, it is also very sensitive to abiotic forcing. The present climate trends of greater precipitation and less intense winters are favoring the vigor of the positive circulation over that of the negative. There are several significant impacts related to these trends. The reduced volume of dense water has impacted the EMED deep circulation, as noted. Less dense-water volume means that its replacement flow ceases before the end of summer, which increases the relative stagnation of the northernmost portion, and favors bottom-water hypoxia. Characteristic of the precipitation trend is more sudden rain and stronger autumnal rainstorms. The former generates strong runoff peaks of high nutrient content that tend to mix into the interior and cause periodic disturbances to the planktonic community. The latter generates an enormous flux of freshwater, some of which mixes or recirculates out of the coastal current and into the interior and subsequently inhibits winter dense-water production and reduces the effectiveness of the winter flushing of eutrophic products.

**Sub-Arctic Sea**

The Arctic Ocean is composed of the mostly ice-covered Polar Sea, consisting of the Eurasian and Canadian Basins, and the mostly ice-free Sub-Arctic Sea, consisting of the basins of the Greenland, Iceland, and Norwegian Seas (Fig. 9). The latter extend southward from the Fram Strait to the Greenland-Scotland Ridge and constitute arguably the most important and complicated thermohaline region of the world’s ocean (cf. Hopkins, 1991). The area produces three dense-water masses that constitute an enormous flux of freshwater, some of which mixes or recirculates out of the coastal current and into the interior and subsequently inhibits winter dense-water production and reduces the effectiveness of the winter flushing of eutrophic products.

![Fig. 9 – The bathymetry of the Arctic Ocean, hatched at depths >2500 m. The lettered features are as follows: CB Canadian Basin, LR Lomonosov Ridge, EB Eurasian Basin, FS Fram Strait, GB Greenland Basin, BS Barents Shelf, MR Mohn Ridge, LB Lofoten Basin, NB Norwegian Basin, IP Icelandic Plateau, DS Denmark Strait, IFR Iceland-Faeroe Ridge, FSC Faeroe-Shetland Channel. The arrows labeled a, b, c refer to the locations of the air-temperature time series of Fig. 10: a) Jan Mayen Island, b) Bear Island, c) Angmagssalik. (From Hopkins, 1991b).](image)
integral part of the thermohaline circulations of both the Polar Sea and the North Atlantic. The thermohaline influence on the Atlantic is a critical aspect of the glacial cycles (cf. Broecker and Denton, 1990). Climate models project the Polar Regions to exhibit the greatest sensitivity to greenhouse warming (e.g. Kattenberg et al. 1996). Significant changes have recently been reported in the Polar Sea, for example, in the water-mass distribution (e.g. Steele and Boyd, 1998) and in decreasing ice cover (Cavalieri et al., 1997). These findings put in question the stability of both the atmospheric and oceanographic feedback loops that control the stability of the Arctic heat balance and thereby the stability of Global Climate.

Fig. 10 – Schematic of the horizontal circulation and location of the three dense-water production areas (upper panel) and the vertical circulations of these products in connection with the North Atlantic and Polar Sea (lower panel).
The Greenland, Iceland and Norwegian Seas all have negative thermohaline circulations and produce dense-water products because of thermal buoyancy losses (cf. Hopkins, 1991a). The water masses and their trajectories are sketched in Figure 10a,b. The actual fresh-water balance for the total basin is slightly positive, but due to the large import of salt and heat from the North Atlantic, only a moderate heat loss is needed to produce dense water. Consequently, the negative thermohaline circulation is very sensitive to atmospheric temperature fluctuations and is critically dependent on a high salinity. In the Greenland Sea, the Greenland Deep Water (GDW) is formed, which drives a deep outflow to the north through Fram Strait into the Polar Sea and an outflow over the Mohn Ridge to the south into the Norwegian basin. There it accumulates, and together with contributions from the intermediate water products, it forms the Norwegian Deep Water (NwDW). The Arctic Intermediate Water (ArIW) forms in the Icelandic Sea where it primarily overflows in the Denmark Strait and over the Iceland-Faeroe Ridge. The large accumulation in the Norwegian Basin mainly overflows through the Faeroe-Shetland Channel and thence moves down-slope and westward being joined by the other two overflows. The result of these combined overflows is the North Atlantic Deep Water (NAtDW), which forms the western bottom boundary current and contributes to the relatively high oxygen and lower carbon-dioxide levels of the deep North Atlantic. Therefore, the thermohaline boundary along the Greenland-Scotland Ridge imposes a strong interactive coupling between the Sub-Arctic’s negative circulation and the North Atlantic. The North Atlantic has a similar coupling with the Mediterranean Sea.

The replacement waters for the Sub-Arctic, required by its overflows to the North Atlantic, enters through Greenland-Scotland-Ridge system on the eastern sides of its passages. Being less dense than the resident waters to the north, the North Atlantic water moves down the pressure gradient and against the Norwegian continental shelf where it forms the large, broad Norwegian Atlantic Current (NwAtC) (Hopkins, 1991a). Atmospheric cooling of the NwAtC forms several intermediate-water products, which we will simply refer to as the Atlantic Intermediate Water (AtIW). The most important AtIW component is the cooled version of the NwAtC that continues northward, mostly through via Fram Strait but also through the Barents Sea, and sinks under the halocline of the Polar Sea. The other products contribute to the dense water accumulations in the Norwegian and Greenland Seas. Hence, the NwAtC provides the main source of heat and salt for the entire Arctic Ocean first by contributing to the surface waters of the Sub-Arctic and then to the sub-halocline waters of the Polar Sea.

However, this is not the entire thermohaline story of the Arctic Ocean. The northern thermohaline boundary with the Polar Sea is more complicated because of the double thermohaline circulation of the Polar Sea. At intermediate and deeper depths, the Polar Sea has a negative circulation, driven by brine-water formation on the shelves and by the deep entrance of the GDW into the Eurasian Basin. This negative circulation couples in Fram Strait with Polar waters flowing south on the west side and Sub-Arctic water masses flowing northward on the east side. In contrast, the surface layer of the Polar Sea has a positive thermohaline circulation for three major reasons. The Polar Sea receives the large runoff of the Eurasian and North American rivers (cf. Aagaard and Carmack, 1989) and it receives freshwater through the one-way flow of lower salinity water through the Bering Strait. On the other hand, it loses very little buoyancy to the atmosphere because the polar ice-cover insulates the underlying waters from surface heat-loss.

This excess surface buoyancy, relative to the adjoining Greenland Sea, generates the positive circulation with its freshwater outflow leaving via the Fram Strait as the East Greenland Current (EGC) and the replacement waters entering via the NwAtC. The EGC carries roughly an equal amount of freshwater in the ice carried southward along the Greenland coast (ibid.) Part of the precarious thermohaline stability of the Arctic system is that this low-salinity water remains geostrophically trapped along the Greenland Coast and during its 1500-km passage it minimally interferes with the negative circulation of the Sub-Arctic (ibid; Hopkins, 1990b). The EGC continues as a coastal current around Greenland, around the Labrador Sea and Baffin Bay and joins the low-salinity surface flow along the East Coast of North America to Cape Hatteras. It has been hypothesized (Wallace, pers. comm.) that the particularly low N:P ratios of the Polar Sea outflow carried by this long trajectory may contribute to the nitrate limitation of primary production along the shelf boundary of the northwestern Atlantic. This would constitute another long-reaching thermohaline, abiotic-biotic interaction.

THERMOHALINE FEEDBACK LOOPS 249
Feedback loops

The ice cover in the Iceland Sea shows a larger variability from year-to-year than on an interannual scale (Malmberg, 1984). The annual air temperatures on the coast of Greenland and from the small islands in the Sub-Arctic show large variations at one-to-three year periodicity (Fig. 11a). In particular, the temperatures from Jan Mayen Island show a damped annual oscillation following the cold spell of the 1960s. Both of these characteristics are suggestive of a self-regulating mechanism. To test the hypothesis that these effects could be influenced by a thermohaline feedback loop, Hopkins (1991b) constructed a simple model of the dense-water production for the Sub-Arctic.

Since each of the three water masses produced have a different water type and residence time in the system, they had to be accounted for separately. The three production areas (Fig. 10 top) were based on the mean SST values from AVHRR images. These three areas summed to the total mean ice-free region of the Sub-Arctic basin. It was assumed that heat lost to the atmosphere by each area contributed either to the preconditioning or to the formation of the associated dense-water product. The dense-water volume began to form when enough heat was removed to bring the temperature equal to that of the target water mass. The entire process was calibrated to mean estimates of atmospheric heat-loss, exchange, and residence times. To complete the atmospheric loop for temperature, deviations in the mean air temperature were calculated from the water temperature changes.

Model simulations of perturbations in the atmospheric temperature and in the amount of ice cover confirmed the feedback-loop processes shown in Figure 12. In another simulation, the air temperatures were forced to match the 1960s cold spell and then left to be influenced by variations in the sea temperature. These results produced a year-to-year variability similar to that observed (Fig. 11b). The co-production of the three different water masses, each having different residence times and sensitivities, complicated the response. In terms of production, the AtIW is the most influenced by air-temperature variability because of its larger area and more direct coupling with the replacement inflow. This supposition is confirmed by the greater amplitudes in the air-temperature oscillations observed at Bear than at Jan Mayen Island (Fig. 11a). The AtIW outflow was responsive to temporal changes in production because of its short (~ 2-yr) residence time. Therefore, during the simulated 1960s cold spell it responded with the largest increased outflow. This responsiveness of the AtIW production to atmospheric temperature variability and its short trajectory through the Denmark Strait suggest that it could generate a corresponding variability in the NAtlDW.
The GDW outflow was not responsive in the model to annual variability in its production because of its large accumulation volume. As explained, its outflow has two trajectories and that to the deep Polar Sea (Eurasian Basin) would be more responsive than that to the Norwegian Basin. In either case, a lessening of its production would eventually show up as a weakening of first the deep Fram Strait exchange and later of the Faeroe-Shetland overflow. The observed, multi-decadal trend in its decreased production, implied by its warmer, saltier, and less oxygenated water type (Aagaard et al. 1991), has probably altered its direct outflow to the Eurasian Basin through Fram Strait. In terms of ice cover, the modeled GDW production was the most sensitive to increased ice cover because it reduced the area of its formation. This relates to the concern of an expanded Polar-Sea freshwater outflow (discussed below) that would either inhibit GDW formation through extended ice cover or through salinity decreases.

Stability

The negative thermohaline circulation of the Sub-Arctic Sea acts as a buffer to short-term climate variability. The buffering action results both from the residence time of the water mass, which decreases the impact of short perturbations, and from the buoyancy compensating effect of the replacement inflow (Fig. 12). The extent to which one could extrapolate this simple depiction to more complicated perturbations and trends is limited for several reasons involving potential interactions between variables and systems. For example, the above exercise served primarily to demonstrate how the thermohaline feedback loop could couple with the atmosphere through the variable of air temperature. While perturbations of the other two variables are also stabilized by a similar process, as per Figure 12, the model formulation did not include external forcing for salinity and ice cover, nor did it include the important interactions between water
temperature and salinity with respect to ice cover, or the control of all three variables by wind-driven circulations and so forth. These would enter in a more complicated model.

The fact that the Sub-Arctic is situated as an intermediary system between the Polar Sea and Atlantic Ocean makes all three thermohaline systems mutually sensitive to changes in a manner that would defeat the stability of any one system. Further, all three of these systems are linked to the atmospheric system of the Northern Hemisphere. Hence, of great concern would be any interactive response to longer-term atmospheric trends that could endanger the stability of the present Arctic ocean-climate system. Several examples follow.

Of most immediate concern is the response of the AtIW because of trends now observed in the surface layer of the Polar Sea that could change the set of feedback loops controlling global climate. Two ongoing trends are being observed in the Polar Sea: increased precipitation (IPCC, 1996) and decreasing ice cover (Cavaliere et al., 1997). Both of these trigger a faster negative circulation for the Polar Sea and therefore a larger replacement flow, i.e. the NwAtC. A concurrent warming trend in the Norwegian Sea could combine with the increased volume flux to give a greater heat flux to the Polar Sea. Evidence of such an increased influx of Atlantic water into the Polar Sea is suggested by the major advance past the Lomonosov Ridge of Atlantic water-mass assemblies (McLaughlin et al., 1996). Another great concern is whether this increased heat input and the decreased ice cover could lead to a partial flip in surface thermohaline circulation such that large portions over the deep Polar Basins would become ice-free and through convection produce great volumes of dense water that would cascade into the Sub-Arctic Basin and into the deep Atlantic. The replacement requirement would draw in much larger quantities of warm Atlantic water and feed the ice-melting process. The dynamical question would be whether convection could break through the surface halocline and whether the Atlantic waters could remain dynamically separated from the waters freshened by Arctic runoff so that they could provide the high salinity source water for the convection process. This is a difficult question, as is also the question of how associated climate feedback loops would stabilize. Both the decreased albedo and the convective process would warm the atmosphere, but would also put much more moisture into the atmosphere, which might then increase the snow cover on the arctic landmasses, again increasing the albedo and the freshwater input to the Polar Sea and thereby its ice cover. If nothing else, these uncertainties underline the vulnerability of the Arctic system and argue for better representations of ocean convective processes and thermohaline feedback loops in climate models.

For the Arctic, a thermohaline circulation switch from positive to negative or vice versa, would be more aptly described as an environmental crisis rather than a “surprise.” The entire Northern Atlantic/Arctic system is presently negative, and active bottom-water production requires the warm subtropical Atlantic waters to move north to replace it. The heat of this surface influx helps warm the atmosphere. The climate reversal of the Younger Dryas is an excellent example of a thermohaline crisis when the North Atlantic switched to a positive circulation. The name refers to the period, circa 10,000 ybp, when glacial conditions returned to the landmasses surrounding the Northern Atlantic. The suggested cause was a shift in the melt-water drainage from the Mississippi to the St Lawrence Rivers causing a flooding of freshwater into the Northern Atlantic and consequently the ice cover to increase to the point of sealing off the dense-water production and thereby the poleward advection of heat (Broecker and Denton, 1990).

Another important aspect of a such mode shift in the Atlantic thermohaline circulation is that stopping production creates a positive feedback loop with respect to the greenhouse effect. The ocean stores roughly fifty times as much carbon as the atmosphere, mostly in the Pacific deep basin because of its size and slow renewal times. The North Pacific has a positive thermohaline circulation with virtually no dense-water production (cf. Warren, 1983). A reversal of the Atlantic thermohaline circulation would stop its advective, poleward heat flux and thereby cool the Northern Hemisphere. On a longer time scale, the lack of deep flushing would allow a larger accumulation of carbon in its deep reservoir and weaken the greenhouse effect and cool the atmosphere. On the contrary, increasing its negative circulation, as in the above scenario of a convecting Polar Sea, would increase the advective heat flux and increase the flushing, which would then bring more carbon dioxide to the surface and accelerate the Greenhouse effect.
CONCLUDING COMMENTS

Thermohaline stability

To assess natural systems it is critically important to understand how they sustain themselves in quasi-stable energy states. Low-entropy, biotic systems increase their resilience against abiotic (external) variability by optimizing how they process fluxes of mass and energy through both internal and external interactions. With many ways to store energy, nutrients, and carbon, they are able to reorganize to accommodate different levels of abiotic input. In contrast, abiotic systems, such as the thermohaline cases we reviewed, are much more limited in how they can reorganize. Thermohaline circulations are driven by heat and salt, and their options for storage are limited by the specific geomorphology of the basin and the circulation of its waters. Generally, their only two modes of reorganization correspond to the positive and negative thermohaline circulations. But for certain dynamics, they can have both modes: in the vertical as the Polar Sea or in the horizontal as the Adriatic Sea.

We can study the range of stability of these systems by considering different geomorphic basins and the different climate regimes that drive them. From a comparative study of different systems we can better understand the thermohaline limits to the stability of a given system. The greater challenge is to understand how the thermohaline structure interacts with the resident biotic system and with other marine systems and the atmospheric.

In the case studies, we have seen how thermohaline feedback loops can stabilize or destabilize both abiotic and biotic components within a system. For example, the stability of the Albemarle Sound is buffered by the intervening Pamlico Sound, such that the salinity of its replacement water is modified by that of the Pamlico exchange with the ocean. The Neuse estuary has a range of thermohaline stability within which the biological production is stabilized and out of which it was destabilized. Atmospheric trends in precipitation, ice cover, and heating have combined to change the buoyancy of the upper layer of the Polar Sea with great potential consequences to the North Atlantic and Arctic systems. Among these would be a reduced ice cover and convection in the Polar Sea. Another would be increased fluxes of fresh-water from the Polar Sea that would inhibit the dense-water production in the Sub-Arctic. Shifts in the buoyancy balances of the Adriatic and Aegean Seas have upset the stability of the Eastern Mediterranean deep waters. It is obvious that when these abiotic examples are connected to the complete suite of biotic interactions in a marine ecosystem, evaluating questions of system stability would be greatly complicated.

Science and the public

A stable, sustainable society depends on a self-regulating mechanism which neither government regulation nor the present market economy can provide (e.g. Henderson, 1999; Hawken et al. 1999; Daly, 1991). The critical missing ingredient is the information on Natural Capital and its value that when inserted into the feedback loop of a sustainable economy would dampen perturbations in resource exploitation and degradation relative to a carrying capacity based on changing population, environmental conditions, and technology.

While the information flow has begun spontaneously, other aspects of the changing role of science will need assistance from both the top (government and corporate support) and the bottom (social and natural scientists). An excellent example of a funding initiative is that of the US/NSF Biocomplexity Program (NSF, <http://www.nsf.gov/home/crssprgm/be>). However, for the science community to have momentum beyond trends in funding will require a basic level of commitment and solidarity from the science community itself. This consensus might be only regarding priority and focus in order that the actual role can evolve. For example, marine research priorities could be split into the categories of 1) data acquisition (control, rates, model input), 2) process studies (as presently, but with greater attention to connections and interactions) and 3) systems approach (evaluations and simulations of systems using all of the above). An emphasis on systems science could be the catalyst for the evolving role for science. It would also provide many challenges connected with simulating natural systems and with developing the methodologies for coupling scientific outputs with economic/social needs.

Some portion of science must have a proactive role regarding planetary sustainability. The ecosystem cannot be managed with political paradigms, and the political, social and economic sectors do not have the methodologies to acquire or interpret data or simulate management questions involving complex systems. Hindsight is useless for irreversible processes. Given the risks of a cascading global
degradation cycle, the science sector must prime the transition with information. And even when information has begun to flow, it is of little value to a public unfamiliar with the concepts needed to put it into a meaningful context. Consider, for example, the slowness with which society is confronting the potential threats pertaining to climate change. A strong element of public education is needed. Fortunately, the awareness curve is exponential because the more one learns the easier it becomes to learn more. That is, of course, provided the learning process is not blocked by contrary conviction.

The need for a holistic assessment on the state of natural systems has a different trajectory than the need of science to understand individual components of natural systems. Governments now need to know the global carrying capacity (although they do not seem to realize this yet). They need to know how much waste we can put in the atmosphere, how many trees we need, how many fish we can take from the sea, and how many children we can have. So how could we organize our approach to answering such questions?

Part of the answer lies in the second time derivative. We cannot assume systems to be in continuous steady state when our objective is to model the system response to variable forcing. In biotic systems, linearity is the exception rather than the rule. Rates of change are insufficient indicators, we need to know how these rates can or will change. We also need more documentation on the ranges of validity of the dynamics we simulate and on the error of omitting interactions. It is the potential interactions and the initial condition that determine the reorganization process. We need to eliminate (or at least try to anticipate) environmental surprises. Another part of the answer lies in complexity. We need to know how a disturbance propagates through a system, how it stimulates reorganization among internal components, and how it changes the function of the system. We also need to consider system memory and the significance of initial condition. We need to know how storage and structure control the response of a system at the onset of an event.

We are really talking here about three steps or questions and a decision: Step 1) Evaluate the function and yield of a system. This is the steady-state question. In anthropogenic terms, it means what does the system do for us as it is and how much yield can we take from it. Step 2) Evaluate deviations from sustainability. This is the trends or first-derivative question. We need to evaluate how the system is degrading (or recuperating) and how this affects the stability of its function and yield. This would require evaluating its interactions with other systems. Step 3) Estimate the probability of irreversible change and its eventual costs. This is the second-derivative or environmental-surprise question. In most cases we can understand the types of changes possible and give some associated probability. Finally, there is the ultimate multidisciplinary decision to be made in the political, social, and economic sphere: 4) How should humans reorganize their interaction with natural systems? How can we utilize technology, through social and economic implementation, to optimize a sustainable use of natural resources?

Each of the first three steps will require a greater understanding of the reorganization processes internal to systems and their interactions with other natural systems and with the anthropogenic system. Steps one and two will require a complex, intelligent monitoring system. Step three will require a complex, evolving network of simulation models. The costs will be a function of the costs to duplicate the necessary function and yield of the system and the collateral costs to interactive systems. How we make the ultimate decision, however, will require a global commitment to achieving sustainability and lots of interactive luck.

ACKNOWLEDGEMENTS

The author gratefully thanks North Carolina State University for its administrative support and the following US and International funding agencies that have provided research support: U.S.A. (NSF, ONR, USCGS, NOAA/Sea Grant, NATO/SA CLANT, EU/Environment, CNR/Italy. My thanks also to Lynn Padgett for her expert editing of the manuscript. And I greatly appreciate the thermal interactions and buoyant support provided by many colleagues, students, friends and family.

REFERENCES


Bethoux, J.P., B. Gentili, J. Raunet and D. Taillez. – 1990. Warm-


