Population, Sustainability, and Earth’s Carrying Capacity

A framework for estimating population sizes and lifestyles that could be sustained without undermining future generations

Gretchen C. Daily and Paul R. Ehrlich

The twentieth century has been marked by a profound historical development: an unwitting evolution of the power to seriously impair human life-support systems. Nuclear weapons represent one source of this power. Yet, even the complexities of global arms control are dwarfed by those inherent in restraining runaway growth of the scale of the human enterprise, the second source of possible disaster. Diminishing the nuclear threat involves relatively few parties, well-established international protocols, alternate strategies that carry easily assessed costs and benefits, short- and long-term incentives that are largely congruent, and widespread recognition of the severity of the threat. In contrast, just the opposite applies to curbing the increasingly devastating impact of the human population. In particular, the most personal life decisions of every inhabitant of the planet are involved and these are controlled by socioeconomic systems in which the incentives for sacrificing the future for the present are often overwhelming.

This article provides a framework for estimating the population sizes and lifestyles that could be sustained without undermining the potential of the planet to support future generations. We also investigate how human activity may increase or reduce Earth’s carrying capacity for Homo sapiens. We first describe the current demographic situation and then examine various biophysical and social dimensions of carrying capacity.

Our analysis is necessarily preliminary and relatively simple; we anticipate that it will undergo revision. Nonetheless, it provides ample basis for policy formulation. Uncertainty about the exact dimensions of future carrying capacity should not constitute an excuse to postpone action. Consider the costs being incurred today of doing so little to halt the population explosion, whose basic dimensions were understood decades ago.

The current population situation

The human population is now so large and growing so rapidly that even popular magazines are referring to the possibility of a “demographic winter” (Time 1991). The current population of 5.5 billion, growing at an annual rate of 1.7%, will add approximately 93 million people this year, equivalent to more than the population of Mexico (unless otherwise noted, demographic statistics are from, or projected from, PRB 1991).

Growth rates vary greatly from region to region. The combined population of less-developed nations (excluding China) is growing at approximately 2.4% annually and will double in 30 years if no changes in fertility or mortality rates occur. The average annual rate of increase in more-developed nations is 0.5%, with an associated doubling time of 137 years. Many of those countries have slowed their population growth to a near halt or have stopped growing altogether.

The regional contrast in age structures is even more striking. The mean fraction of the population under 15 years of age in more-developed countries is 21%. In less-developed countries (excluding China) it is 39%; in Kenya it is fully 50%. Age structures so heavily skewed toward young people generate tremendous demographic momentum. For example, suppose the total fertility rate (average completed family size) of India plummeted over the next 33 years from 3.9 to 2.2 children (replacement fertility). Under that optimistic scenario (assuming no rise in death rates), India’s population, today some 870 million, would continue to grow until near the end of the next century, topping out at approximately 2 billion people.

The slow progress in reducing fertility in recent years is reflected in the repeated upward revisions of United Nations projections (UNFPA 1991). The current estimate for the 2025...
population is 8.5 billion, with growth eventually leveling off at approximately 11.6 billion around 2150. These projections are based on optimistic assumptions of continued declines in population growth rates.

Despite the tremendous uncertainty inherent in any population projections, it is clear that in the next century Earth will be faced with having to support at least twice its current human population. Whether the life-support systems of the planet can sustain the impact of so many people is not at all certain.

Environmental impact

One measure of the impact of the global population is the fraction of the terrestrial net primary productivity (the basic energy supply of all terrestrial animals) directly consumed, co-opted, or eliminated by human activity. This figure has reached approximately 40% (Vitousek et al. 1986). Projected increases in population alone could double this level of exploitation, causing the demise of many ecosystems on whose services human beings depend.

The impact (I) of any population can be expressed as a product of three characteristics: the population's size (P), its affluence or per-capita consumption (A), and the environmental damage (T) inflicted by the technologies used to supply each unit of consumption (Ehrlich and Ehrlich 1990, Ehrlich and Holdren 1971, Holdren and Ehrlich 1974).

\[ I = PAT \]

These factors are not independent. For example, T varies as a nonlinear function of P, A, and rates of change in both of these. This dependence is evident in the influence of population density and economic activity on the choice of local and regional energy-supply technologies (Holdren 1991a) and on land management practices. Per-capita impact is generally higher in very poor as well as in affluent societies.

Demographic statistics give a misleading impression of the population problem because of the vast regional differences in impact. Although less-developed nations contain almost four-fifths of the world's population and are growing very rapidly, high per-capita rates of consumption and the large-scale use of environmentally damaging technologies greatly magnify the impact of industrialized countries.

Because of the difficulty in estimating the A and T factors in isolation, per-capita energy use is sometimes employed as an imperfect surrogate for their product. Using that crude measure, and dividing the rich and poor nations at a per-capita gross national product of $4000 (1990 dollars), each inhabitant of the former does roughly 7.5 times more damage to Earth's life-support systems than does an inhabitant of the latter (Holdren 1991a). At the extremes, the impact of a typical person in a desperately poor country is roughly a third to one of that of an average citizen of the United States. The US population has a larger impact than that of any other nation in the world (Ehrlich and Ehrlich 1991, Holdren 1991a,b).

The population projections and estimates of total and relative impact bring into sharp focus a question that should be the concern of every biologist: if not every human being; how many people can the planet support in the long run?

The concept of carrying capacity

Ecologists define carrying capacity as the maximal population size of a given species that an area can support without reducing its ability to support the same species in the future. Specifically, it is "a measure of the amount of renewable resources in the environment in units of the number of organisms these resources can support" (Roughgarden 1979, p. 305) and is specified as K in the biological literature.

Carrying capacity is a function of characteristics of both the area and the organism. A larger or richer area will, ceteris paribus, have a higher carrying capacity. Similarly, a given area will be able to support a larger population of a species with relatively low energetic requirements (e.g., lizards) than one at the same trophic level with high energetic requirements (e.g., birds of the same individual body mass as the lizards). The carrying capacity of an area with constant size and richness would be expected to change only as fast as organisms evolve different resource requirements. Though the concept is clear, carrying capacity is usually difficult to estimate.

For human beings, the matter is complicated by two factors: substantial individual differences in types and quantities of resources consumed and rapid cultural (including technological) evolution of the types and quantities of resources supplying each unit of consumption. Thus, carrying capacity varies markedly with culture and level of economic development.

We therefore distinguish between biophysical carrying capacity, the maximal population size that could be sustained biophysically under given technological capabilities, and social carrying capacities, the maxima that could be sustained under various social systems (and, especially, the associated patterns of resource consumption). At any level of technological development, social carrying capacities are necessarily less than biophysical carrying capacity, because the latter implies a human factory-farm lifestyle that would be not only universally undesirable but also unattainable because of inefficiencies inherent in social resource distribution systems (Hardin 1986). Human ingenuity has enabled dramatic increases in both biophysical and social carrying capacities for H. sapiens, and potential exists for further increases.

Carrying capacity today. Given current technologies, levels of consumption, and socioeconomic organization, has ingenuity made today's population sustainable? The answer to this question is clearly no, by a simple standard. The current population of 5.5 billion is being maintained only through the exhaustion and dispersion of one-time inheritance of natural capital (Ehrlich and Ehrlich 1990), including topsoil, groundwater, and biodiversity. The rapid depletion of these essential resources, coupled with a worldwide degradation of land (Jacobs 1991, Myers 1984, Postel 1989) and atmospheric quality (Jones and Wigley 1989, Schneider 1990), indicate that the human enterprise has not only exceeded its current social carrying capacity, but it is actually reducing future potential biophysical carrying capacities by depleting es-
sential natural capital stocks. 1

The usual consequence for an animal population that exceeds its local biophysical carrying capacity is a population decline, brought about by a combination of increased mortality, reduced fecundity, and emigration where possible (Klein 1968, Mech 1966, Scheffer 1951). A classic example is that of 29 reindeer introduced to St. Matthew Island, which propagated to 6000, destroyed their resource base, and declined to fewer than 50 individuals (Klein 1968). Can human beings lower their per-capita impact at a rate sufficiently high to counterbalance their explosive increases in population?

Carrying capacity for saints. Two general assertions could support a claim that today’s overshoot of social carrying capacity is temporary. The first is that people will alter their lifestyles (lower consumption, A in the I = PAT equation) and thereby reduce their impact. Although we strongly encourage such changes in lifestyle, we believe the development of policies to bring the population to (or below) social carrying capacity requires defining human beings as the animals now in existence. Planning a world for highly cooperative, antimaterialistic, ecologically sensitive vegetarians would be of little value in correcting today’s situation. Indeed, a statement by demographer Nathan Keyfitz (1991) puts into perspective the view that behavioral changes will keep H. sapiens below social carrying capacity:

If we have one point of empirically backed knowledge, it is that bad policies are widespread and persistent. Social science has to take account of them [our emphasis].

In short, it seems prudent to evaluate the problem of sustainability for selfish, myopic people who are poorly organized politically, socially, and economically.

Technological optimism. The second assertion is that technological advances will sufficiently lower per-capita impacts through reductions in T that no major changes in lifestyle will be necessary. This assertion represents a level of optimism held primarily by nonscientists. (A 1992 joint statement by the US National Academy of Sciences and the British Royal Society expresses a distinct lack of such optimism). Technical progress will undoubtedly lead to efficiency improvements, resource substitutions, and other innovations that are currently unimaginable. Different estimates of future rates of technical progress are the crux of much of the disagreement between ecologists and economists regarding the state of the world. Nonetheless, the costs of planning development under incorrect assumptions are much higher with overestimates of such rates than with underestimates (Costanza 1989).

A few simple calculations show why we believe it imprudent to count on technological innovation to reduce the scale of future human activities to remain within carrying capacity. Employing energy use as an imperfect surrogate for per-capita impact, in 1990 1.2 billion rich people were using an average of 7.5 kilowatts (kW) per person, for a total energy use of 9.0 terawatts (TW; 10^{12} watts). In contrast, 4.1 billion poor people were using 1 kW per person, and 4.1 TW in aggregate (Holdren 1991a). The total environmental impact was thus 13.1 TW.

Suppose that human population growth were eventually halted at 12 billion people and that development succeeded in raising global per capita energy use to 7.5 kW (approximately 4 kW below current US use). Then, total impact would be 90 TW. Because there is mounting evidence that 13.1 TW usage is too large for Earth to sustain, one needs little imagination to picture the environmental results of energy expenditures some sevenfold greater. Neither physicists nor ecologists are sanguine about improving technological performance sevenfold in the time available.

There is, indeed, little justification for counting on technological miracles to accomodate the billions more people soon to crowd the planet when the vast majority of the current population subsists under conditions that no one reading this article would voluntarily accept. Past expectations of the rate of development and penetration of improved technologies have not been fulfilled. In the 1960s, for example, it was widely claimed that technological advances, such as nuclear agroindustrial complexes (e.g., ORNL 1968), would provide 5.5 billion people with food, health care, education, and opportunity. Although the Green Revolution did increase food production more rapidly than some pessimists (e.g., Paddock and Paddock 1967) predicted, the gains were not generally made on a sustainable basis and are thus unlikely to continue (Ehrlich et al. 1992). At present, approximately a billion people do not obtain enough dietary energy to carry out normal work activities.

Furthermore, as many nonscientists fail to grasp, technological achievements cannot make biophysical carrying capacity infinite. Consider food production, for example. Soil can be made more productive by adding nutrients and irrigation; yields could possibly be increased further if it were economically feasible to grow crops hydroponically and sunlight were supplemented by artificial light. However, biophysical limits would be reached by the maximal possible photosynthetic efficiency. Even if a method were found to manufacture carbohydrates that was more efficient than photosynthesis, that efficiency, too, would have a maximum. The bottom line is that the laws of thermodynamics inevitably limit biophysical carrying capacity (Fremlin 1964) if shortages of inputs or ecological collapse do not intervene first.

Sustainability

A sustainable process is one that can be maintained without interruption, weakening, or loss of valued qualities. Sustainability is a necessary and sufficient condition for a population to be at or below any carrying capacity. Sustainable development has thus been defined as “development that meets the needs and aspirations of the present without compromising the ability of future generations to meet their own needs” (Brundtland 1987, p. 43). Implicit in the desire for sustainability is the moral conviction that the current generation should pass on its inheritance of natural wealth, not unchanged, but diminished in potential to support future generations.
In any discussion of sustainability, it is clearly necessary to establish relevant temporal and spatial scales. The time scale that will be considered here is tens of human generations—that is, hundreds of years to a millennium. The spatial scale is obviously constrained by the size of the planet, a closed system for most purposes. Though trade enables populations to sustainably exceed local and regional carrying capacities, all accounts must balance for Earth as a whole.

**Classification of resources.** How does one determine a sustainable level of consumption? To address this question, we start by specifying several resource types and analyzing the constraints on their use independently. Then, the paramount importance of interactions deriving from the simultaneous use of a resource required for multiple activities is considered. We also highlight throughout means by which humanity could increase the maximum sustainable levels of resource consumption (dimensions of biophysical carrying capacity).

Our scheme involves the somewhat arbitrary classification of continuously distributed elements into discrete units to bring into focus key aspects of sustainability. First, there are the resources that provide free services to humanity without necessarily undergoing depletion or degradation (Table 1, first column). These resources include microbial nutrient cyclers and soil generators, natural pest-control agents, and pollinators of crops. Of special importance are the forests, which help to maintain a balance of gases in the atmosphere, to ameliorate local climate, to provide habitat for wildlife, to control erosion, and to run the hydrologic cycle. Other resources, such as food, drinking water, energy, and the capacity of the environment to absorb pollutants, are necessarily consumed, dispersed, or degraded as the benefits are derived from them.

Second, there is an important distinction in practice between renewable and nonrenewable resources, although renewal rates are continuously distributed. Renewable resources tend to be flow-limited and are reconstituted after human consumption or dispersion through natural processes driven by solar energy (which may be enhanced by human investment, as when trees are planted). Nonrenewable resources are generally stock-limited and have either very low or no renewal rates and prohibitive reconstitution costs (though one or more recyclings before ultimate discard may be possible; Ehrlich et al. 1977). The rate of degradation and erosion of topsoil (according to one estimate a net 25 billion tons erosion loss per year; Brown and Wolf 1984) is so much in excess of its rate of creation that soil has been turned into an essentially nonrenewable resource on any relevant time scale. The same can be said of groundwater in many aquifers (e.g., Wittwer 1989) and biodiversity (Ehrlich and Wilson 1991).

Last, resources may be further classified into two types: those for which substitutes are either currently or foreseeably available (substitutable resources) and those for which complete substitution at the required scale is currently and foreseeably impossible (essential resources). Substitutable resources include fossil fuels, some metals and minerals, and some natural fibers. Essential resources include fertile soils, fresh water, and biodiversity. The classification of some resources may vary depending on the manner in which they are used; for example, forests as sources of wood are substitutable resources because wood has substitutes for most purposes, whereas forests as sources of ecosystem services generally constitute essential resources.

Maximum sustainable use. The maximum sustainable level of use (MSU) of a resource depends on how it is classified with respect to the preceding attributes and on socioeconomic factors. Using the classification scheme of Table 1, we may now specify a theoretical MSU for each resource type, which represents dimensions of biophysical carrying capacity.

In the case of resources from which humanity may benefit without causing their depletion or degradation, MSU is proportional to the total extent of the resource: the greater the forested area, for example, the greater the scale of ecosystem services provided by forest. In this general case, sustaining maximal use is a matter of safeguarding the ability of such resources to provide humanity with services. For uses that necessarily alter resources in the process of deriving benefits from them, MSU depends on the resource's renewability and substitutability.

Let us consider nonrenewables first. No resources that are absolutely essential for human life have been classically considered nonrenewable except those for which supplies are so large (e.g., calcium) as to make worrying about them pointless. A notable exception may be time or opportunities to prevent irreversible, possibly catastrophic consequences of anthropogenic impacts on the environment.

Numerous nonrenewable substitutable resources are critical to maintaining certain features of today's civilization, although their disappearance

---

### Table 1. Resource classification scheme with some examples. This classification scheme makes explicit three key parameters that determine the nature of the maximum sustainable level of use for a particular resource. The examples provided are by no means exhaustive.

<table>
<thead>
<tr>
<th>Resource type (at current use rates)</th>
<th>Essential</th>
<th>Substitutable</th>
<th>Time or opportunity</th>
<th>Nonrenewable energy sources (e.g., fossil fuels), some other minerals</th>
<th>Solar energy, fresh water, some soil used for agriculture</th>
<th>Wood for construction, any particular food type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonrenewable</td>
<td>Stratospheric ozone, tropical forests, biodiversity</td>
<td>Materials that supply some services (e.g., diamonds and gold for aesthetic and wealth repository purposes)</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Renewable</td>
<td>Ecosystem elements that supply services (e.g., soil microbes, some temperate forests, pollinators)</td>
<td>Species that supply some services (e.g., animals for power, transport, insulin, and vaccines; trees for cooling buildings)</td>
<td>None</td>
<td>None</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

---

3See footnote 1.
would not threaten human existence. Iron, for example, is used heavily in the production and transport of energy and goods in industrial societies. By definition, there is no sustainable rate of consumption of non-renewables; the closest approximation is a quasisustainable consumption rate equivalent to (or lower than) the rate of generation of substitutes. The primary difficulty in the use of nonrenewables is not exhaustion per se (because quantities are generally gigantic), but rather the technical, economic, environmental, and sociopolitical difficulties associated with declining quality of the resources (with respect to, for example, distance, depth, and concentration) and with the transitions to substitutes (e.g., Holden 1991a).

At first glance, it might seem that stocks and flows of renewable resources would require the least effort to maintain simply because they are regenerateable for us. However, increasing human demands on the biophysical environment make it difficult to limit the use of many renewable resources to a sustainable rate. It is therefore critical to consider how MSUs of renewable resources vary as a function of those stocks, that is, how human activity may increase or reduce those elements of biophysical carrying capacity.

For a renewable essential resource that is necessarily consumed, degraded, or dispersed in the extraction of value from it, the MSU is equivalent to its renewal rate. MSU (and maximum sustainable yield) increases monotonically with the global extent of resource stocks (e.g., agricultural soils, harvested forest, and groundwater) above a critical point. As the land area covered with productive agricultural soil, supporting intact forest, or underlain by freshwater aquifers is reduced, the MSU of these resources is proportionally diminished. The minimum represents the point below which the constituent stocks are so small that the resource cannot be used sustainably. For example, very thin soils are agriculturally unproductive (UNEP 1984), and regeneration of trees may fail in small remnant forest patches subject to deleterious edge and isolation effects.

Interestingly, surface water also features a linear relationship between MSU and stock, and it illustrates a case where humanity may increase MSU by altering the spatial and temporal distribution of the resource. Although humanity exercises substantial control over the distribution of water among different (natural or artificial) channels and reservoirs (White 1988), it has relatively little direct control of the total stock. Furthermore, silting of dams and salinization of agricultural water may represent barriers to increasing the long-term MSU of water through anthropogenic manipulation. Recently, humanity has unwittingly reduced the total annual input to some surface water systems through deforestation and desertification (Myers 1989). More dramatic changes in regional stocks of surface water are expected as a consequence of global warming (Gleick 1989, Schneider 1990, Tegart et al. 1990).

The extraction of resources is generally managed not at the global scale but at local or regional levels. Several functional relationships between MSU and a single local resource stock are possible. The curve in Figure 1a describes a general relationship between MSU of agricultural soil and the stock (soil depth). While soil depth remains significantly greater than the rooting depth of crops or other plants, soil loss has little or no negative effect on productivity, but productivity decreases with soil depth below this threshold. Initially negligible costs of losing soil to erosion may become steep as soil thins below this threshold (called the critical point, C*). The soil depth on most of the cropland in Haiti appears to be substantially below C* (Terborgh 1989, WRI 1992a). Agricultural productivity worldwide is suffering because of such land degradation (UNEP 1984).

The local depletion of aquifers also exemplifies this general relationship between a single stock of a renewable resource and its MSU (Figure 1b). MSU is equivalent to the rate of recharge at any stock above C*. MSU is constant across nearly all values of stock because the renewal rate is largely stock independent. At stock levels below C*, aquifers may suffer from salinization or collapse (Dunne and Leopold 1978), reducing MSU.

There are two important differences between the management of soils and aquifers, although the func-
tional forms below the critical point are uncertain. First, many aquifers contain orders of magnitude more water than the critical volume, whereas soils are rarely more than a few times deeper than the critical depth. Second, MSU of water from aquifers may decline more rapidly below C* than that of many soils (NAS 1989).

A hypothetical relationship between MSU and a forest harvestable at maximum sustainable yield is depicted in Figure 2. Though the precise functional form depends on forest type and harvesting method, the rate of forest regeneration is highest at a biomass density below the maximum attainable. At extremely high densities, trees suffer from overcrowding; at very low densities, microclimatic and other conditions may become unfavorable for germination and sapling recruitment.

Where resources in high demand and in short supply are overharvested, a positive feedback cycle is established, thereby sequentially depleting the stocks and lowering the MSUs. For example, overharvesting of fuelwood, the primary source of energy for more than half of the world’s population, has created severe local and regional shortages. To supply domestic energy, these shortages are countered by overharvesting increasingly distant supplies and by burning animal dung and crop residues, important inputs to the maintenance of soil productivity (WRI 1992b). For any essential resources that may limit the size of the human population (e.g., fertile soil, forest products and services, and fresh water), depletions constitute a reduction in biophysical carrying capacity of the planet.

MSUs of renewable substitutable resources that are necessarily consumed, degraded, or dispersed are also equivalent to their renewal rates (that may be enhanced by human investment). Maintenance of the function served by such resources could also be sustained if the supply were exhausted at a rate less than or, at most, equivalent to the rate of generation of substitutes. Thus, coal and then petroleum and other substitutes replaced wood as a primary source of industrial energy.

Maximum sustainable abuse. We next consider the passive use of natural biogeochemical processes to absorb waste and to reconstitute component resources therein, also elements of biophysical carrying capacity. This analysis on sustaining output rates complements the foregoing one that concerns sustaining input rates. The maximal sustainable emission rate of a pollutant enters the environment (maximum sustainable level of abuse; MSA) is defined as the rate above which unacceptable damage is caused. Specifying levels of damage that are unacceptable is a subject of a complex literature on risk analysis (see for example, Ehrlich and Ehrlich 1991, Kates et al. 1985).

Humanity exercises some control over four parameters relating to MSA: the type of pollutant released, the spatial distribution of the pollutant, the total stock of pollutant in the environment, and the scale and health of natural (or human-made) ecosystems that are meant to absorb the pollutant. In this article, we explore the following two relationships: first, that between MSA and the scale and health of the ecosystem(s) into which the waste is released; and, second, that between the total stock of accumulated pollutant and the ability of the environment to buffer *H. sapiens* from harmful effects.

Pollutants whose rates of removal are limited, at least in part, by biological processes differ from those whose removal rates are not biolimited. Removal may be achieved by degradation into benign products, dilution to harmless levels, or transfer into sinks. Virtually all organic wastes (e.g., sewage and pulp mill effluents) are biolimited. Examples of pollutants whose removal rates are not biolimited include asbestos and radioactive materials.

MSA is a function of the pollutant’s distribution and rate of removal and of the sensitivity of the affected systems to its concentration. For a given spatiotemporal distribution of pollutant, MSA is the level of emission that produces the highest concentration of pollutant that can be tolerated by the most sensitive system element. If the removal mechanism is the most sensitive, then MSA is equivalent to the maximal sustainable average rate of removal. For example, MSA for organic waste flushed into an aquatic system is equal to the maximal emission rate that does not lead to eutrophication. System elements other than those involved in removal may be more sensitive. Thus, for a toxic waste that can be degraded by specialized bacteria, MSA may be limited by the sensitivity of components of the recipient ecosystems other than the bacteria (e.g., shellfish, fishes, seabirds, and marine mammals in the case of oil spilled into the oceans). Variation in the emission or removal rates must be incorporated into the calculation of MSA. Although the average removal rate may be sufficient to prevent long-term buildup of a pollutant, variation in the rate may allow temporary but harmful concentrations to develop, as in the cases of air pollution in city basins that are only periodically swept clean by winds or of acid pulses associated with the spring melt of acid snow.

MSA may be increased in two ways. The first is by manipulation of the distribution of pollutant into concentrations that maximize the removal rate of buffering capacity of the environment. The second is by enhancement of the removal rate by increasing the extent and capacity of systems involved in its removal, be they natural ecosystems or sewage treatment plants.

The same analysis applies to pollutants whose rates of degradation or uptake by sinks are not biolimited. Although their removal rates are independent of the scale and capacity of ecosystems, their MSAs may depend on these factors to the extent that ecosystems buffer humanity and other life-forms from negative impacts by, for instance, dilution. Any level of waste generation could be considered quasisustainable (even for

Figure 2. The general relationship between the maximum sustainable level of use (MSU) of a forest under harvest and its biomass density.

BioScience Vol. 42 No. 10

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
pollutants with no degradation rates, such as asbestos) until the capacity of the environment to buffer human and its life support systems from unacceptably harmful effects is exceeded.

Interactions. The preceding analysis enables calculation of upper bounds on carrying capacities by dividing each MSU and MSA by the minimal or desired average per-capita use of abuse and finding the minimum among all those resources. However, the simultaneous use of different resources usually involves complex, indirect interactions that constrain MSUs and MSAs of a resource required for multiple activities (e.g., forests).

A systems approach is required to keep account of how one activity may impinge on another. To determine a sustainable use of coal, for example, one must account for the damage (e.g., in the form of acid precipitation, strip mining, and global warming) done to natural systems that reduces MSUs and MSAs of those systems. Sustainable farming requires similar comparison of all marginal costs (including decreases in MSAs of soils and water supplies) of applying pesticides and fertilizers to the marginal benefits derived in short-term increases in yield.

Furthermore, a given activity may cause perturbations that have unintended, indirect effects on other system elements. In the case of marine systems, for example, the MSU of a harvested species may depend not only on its own population dynamics (stock-dependent renewal rate), but on the importance of that species in controlling the population dynamics of other species. Harvesting high on the food chain may trigger undesirable population explosions of species lower down. Similarly, harvesting organisms low on the food chain (e.g., krill) may result in the collapse of populations of valued species that consume them (Orians 1990).

The resolution of conflicting demands on interdependent resources involves a complex set of social and economic considerations. Biologists can contribute by describing quantitatively alternative patterns of sustainable use and the relative magnitudes of the carrying capacities resulting from each.

Lag times. A crucial difficulty in assessing whether a given human activity is sustainable is the time that passes between the onset of the activity and human perception of its impact. A delay in perceiving the impact may result from either an actual lag time before its manifestation or from an inability to detect the impact under routine monitoring.

In the case of CFC-catalyzed ozone depletion, there is an actual lag time of approximately a decade between the release of an average CFC molecule and its arrival to the upper atmosphere where it is active. Yet, ozone thinning was only predicted and then detected approximately half a century after freons first came into use. The delay between predicting (Arrhenius 1896) and detecting global warming with certainty is apparently more than a century (Tegart et al. 1990, Schneider 1990); by the time the effects are manifest, irreversible deleterious changes may have occurred (Daily et al. 1991).

Social dimensions of carrying capacity
Social dimensions of carrying capacity include lifestyle aspirations, epidemiological factors, patterns of socially controlled resource distribution, the disparity between private and social costs, the difficulty in formulating rational policy in the face of uncertainty, and various other features of human sociopolitical and economic organization. Although the full complexity of such social dimensions requires investigation beyond the scope of this article, as illustrations, we briefly outline some of the issues surrounding discounting, the global commons, international trade, and prices.

Discounting over time. There are numerous situations (sometimes called social traps), in which the immediate, local incentives are inconsistent with the long-run, global best interest of both the individual and society, and with the maintenance of carrying capacity (Costanza 1987, Cross and Geyer 1980, Platt 1973). One of the most pervasive causes of social traps is the natural human tendency to discount costs that appear remote, either in time or space.

The most straightforward reason for discounting is to adjust for the time value of money: the value of $1000 delivered today is higher than that of $1000 to be delivered in ten years because of benefits that can be derived from investing the money over the decade. Discounting is done routinely in the context of cost/benefit analysis and has enormous influence on fiscal policy in every arena (e.g., Lind 1982).

Although, in principle, discounting is valid, two problems make discounting over a substantial time horizon (several decades or more) a gamble with the welfare of future generations. Estimates of future costs and benefits are uncertain, and there is both subjectivity and uncertainty in the selection of an appropriate discount rate.

Economists have great difficulty assigning monetary value to many of today's environmental amenities (e.g., clean air and national parks) and risks (e.g., global warming and ozone depletion), much less those of the future. When future costs are certain, a risk-averse policy would require discounting less than if they could be predicted with certainty. However, when analysts cannot agree on the uncertainties, too often they make no adjustment at all in the discount rate. The result is an underestimate of potential future costs, such that projects that imperil future generations appear more favorable than they should. These uncertainties are compounded over the period for which the calculation is made; the longer the time horizon, the greater the gamble. And when essential resources are involved, that gamble is with future carrying capacities.

The problem with discounting is not simply that decision makers often fail to apply it appropriately. The very process of discounting (especially at rates as high as 10%) encourages the public to underestimate the importance of future costs and defer their payment. Consider the problem of determining whether society would profit by taking measures now to deter the onset of global warming. Suppose that inaction will result in a known and certain cost of $100 billion to be incurred in 100 years. Discounted at 10% (on an annual basis), the present value of that cost is reduced to a mere $7.2 million. In a cost/benefit framework, investment in any deterrent

November 1992 767

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
whose net immediate cost exceeded $7.2 million seems irrational. But that discounted cost is so deceptively small that society may foolishly fail to invest even that minimal amount to solve a potentially serious future problem. Choosing not to take action now presumes that posterity will be richer than we are, easily able to pay the $100 billion. In the recent past, successive generations have indeed enjoyed ever greater average wealth, but this trend may not continue until the time comes to pay for these deferred costs (Lind 1982; see also, for example, Fuchs and Rekis 1992). In short, this method of analysis should not be applied to long-term resource management because it constitutes a recipe for a growing burden of environmental debt, resulting in lower future carrying capacities.

**Discounting by distance.** Another form of discounting, also important and innate in policy judgements relevant to carrying capacity, is discounting over distance. The significance of events (including the magnitude of benefits and costs) occurring at a distance is discounted. The distance may be measured in strictly geographic terms, or it may be remoteness in a social, economic, or political sense. Discounting over distance is reflected in several dimensions of human behavior and judgement. Consider how societies value domestic environmental health relative to that abroad. Japan is using timber stripped from virgin forests in several nations (including the United States) for low-quality products such as concrete forms, while carefully protecting its own forests. Twenty-five percent of all pesticides exported from the United States are heavily restricted or banned by the United States and other industrialized nations (Weir and Schapiro 1981). The German government made little effort to control industrial emissions until the effects of acid precipitation were manifest in its own forests and soils (to the tune of costing $1.4 billion per year). By then, approximately 18,000 Swedish lakes had acidified to the point that fish stocks were severely reduced, in part due to German emissions (Myers 1984).

In some instances, discounting by distance is clearly in the best interest of the discounters, but misjudgement of the relevant distance may exact a penalty. Overestimation of distance contributes to the extraction and sale, at below-market values, of natural resources (such as timber) from regions that are geographically and socioeconomically remote from policy centers in Washington, DC (e.g., Alaska and Colorado), and clearly confers a net cost to the United States (Wirth and Heinz 1991). Overestimates of the relevant distance have led to profound environmental problems with direct implications for carrying capacity. For example, until recently, the upper atmosphere was considered so remote as to encourage emission of airborne pollutants that did not cause local or regional smog problems. It came as a surprise that the connections between the gaseous composition of the seemingly distant stratosphere and our day-to-day lives are actually very tight (Daily et al 1991). Similarly, the ability of humanity to vastly alter global biogeochemical cycles through local and regional habitat conversion has only become apparent in recent decades.

Currently, the many indications that human society has exceeded social carrying capacity and is paying a price for it are barely noticed. The negative impact of human activity on the planet usually manifests itself first to those whose lives are tightly dependent on the health of fragile, local ecosystems. Yet, by the time many current environmental problems directly affect decision-makers, whose lives are buffered by distance and economic well-being, it will be far too late to correct them. Ecologist Thomas Lovejoy’s program of taking policy-makers and celebrities to tropical forests has helped make apparent the intimate connections to parts of the biosphere that are often misperceived as remote.

For different reasons, discounting over time and distance both encourage behavior that may reduce carrying capacity for future generations. Pressing economic problems often cause developing nations to apply higher discount rates to the future cost of depleting essential resources (as in accepting toxic wastes and environmentally damaging industries rejected by rich countries). Discounting over distance fosters the illusion that wealthy nations and individuals can afford to ignore the increasingly desperate plight of the poor.

**The global commons.** There are several reasons why it is in the selfish best interest of developed nations to narrow the gap between rich and poor. First, it will help the developing nations to protect their vast reservoirs of biodiversity, whose destruction affects at least two major elements of carrying capacity. The need for wild plants and microorganisms, which already supply the active ingredients in more than 25% of modern pharmaceuticals, may become acute as the human population grows more susceptible to disease (Ehrlich and Ehrlich 1990). Biodiversity is also critical to maintaining crop resistance to pests and drought, supplying the raw materials for genetic engineering and thus hopefully permitting the future phenomenal boost in agricultural yields required to feed an exponentially growing population (Ehrlich et al. 1992).

Second, developing nations have the power to degrade severely the entire planet’s life support systems simply by following development paths taken by the rich. Elementary calculations indicate that the mobilization of coal reserves (e.g., in China or India) to fuel even a modest increment of development could overwhelm any efforts by industrialized nations to compensate by reducing their own greenhouse gas emissions (Ehrlich and Ehrlich 1989). Similarly, large increases in methane and nitrous oxide fluxes would accompany planned expansion of agriculture and the continued destruction of tropical forest. The rapid deployment of less-damaging technologies (such as solar-hydrogen energy technologies) in developed nations and their transfer to the rest of the world is required to secure just this atmospheric element of the global commons.

Third, the ever-growing disparity between rich and poor carries forbidding implications for social carrying capacity, including intensifying economic dislocation and social strife as the transfer of capital, labor, and refugees across steepening gradients accelerates. Political challenges also loom large as the ranks of those with little to lose increase, nuclear capability pro-
siderates in the developing world, and vulnerability to terrorism increases (e.g., Schneider and Mesirow 1976).

In short, there is no lifeboat escape possibility for the rich. All nations will have to come to grips with the limits to carrying capacity. Unless measures are taken by the rich to facilitate sustainable development, the continued destruction of humanity’s life support systems (and a reduction in biophysical carrying capacity) is virtually guaranteed.

**International trade.** Trade may increase global biophysical carrying capacity by lifting regional constraints arising from the naturally heterogeneous distribution of resources. If there were no trade at all, then global biophysical carrying capacity would equal the sum of all local biophysical carrying capacities. Trade may also increase global biophysical carrying capacity through the increased efficiency that results from regional specialization in the production of goods.

Exceeding local and regional carrying capacities on a sustainable basis through trade has the unfortunate effect of encouraging the “Netherlands fallacy” (Ehrlich and Holdren 1971): the idea that all regions could simultaneously sustain populations that sum to more than global carrying capacity. Regional and local development plans need to account for the global balance of trade in resources.

The optimal size of resource catchment areas needs consideration with respect to economies of scale and the incentives for sustainable resource management. Empirical evidence suggests that economic incentives favor better management of natural resources by local communities with long-term stakes in sustainability than by distant parties driven to maximize short-term profit (see examples in Ehrlich and Ehrlich 1991). A better understanding is needed of the tradeoffs between the efficiency associated with large industries and the better quality of local resource management.

Finally, the organization and regulation of international commerce is extremely important to evaluation of carrying capacity, but it is also complex and poorly understood (see, e.g., Culbertson 1991, Daly and Cobb 1989, Keynes 1933). For example, standard economic thought tends to support free trade. However, completely unregulated international trade could reduce carrying capacity by tending to diminish international diversity, thereby increasing the vulnerability of nations to disasters in other regions (e.g., droughts in distant grain belts) and limiting their ability to learn lessons from their own successes and failures (e.g., Culbertson 1991).

**Prices.** Prices relate to both biophysical and social carrying capacities in at least two important ways. First, underpricing of resources encourages unsustainable management. Underpricing often occurs because future generations have no means of making their demands for a resource known. The future demand for the water in the Ogallala aquifer clearly is not reflected in its current price. One solution would be to regulate prices of essential resources to keep their use sustainable.

Prices also play an important role in the rates of innovation. High prices constitute incentives for research and development of technologies that are more efficient or that substitute more abundantly for scarce resources. Such price-induced innovation appears to be the rule and can be seen clearly in the development of agriculture (Hayami and Ruttan 1985). The price of food is obviously related to the agricultural dimension of biophysical and social carrying capacities.

**Achieving sustainability**

We wish to reemphasize that our analyses are necessarily preliminary, intended to provide a framework for subsequent more-detailed and quantitative studies. In particular, central determinants of social carrying capacity lie in the domain of interactions among resources, among sociopolitical and economic factors, and between biophysical and social constraints. However, the complexity of these interactions makes it unlikely that they will be sufficiently well evaluated in the next several decades to allow firm calculations of any carrying capacity. From a policy perspective, the current great uncertainty in future social carrying capacity is irrelevant because the human population is likely to remain above that carrying capacity for decades at least.

Global assessments of MSUs and MSAs of critical resources such as forests and the atmosphere should be undertaken immediately, in the tradition already established for greenhouse gases. Such assessments would provide measures of relative contributions of nations to the preservation or destruction of the global commons. They could thus form the basis for international treaties and possible control schemes, such as the issuing of tradable permits for consumption of fractions of global MSUs and MSAs.

Nations and regions should evaluate MSUs and MSAs for their key resources. Even cursory examinations can be informative (e.g., Daly 1990). Fresh water, both surface and underground, is an obvious top candidate for evaluation in many regions, including the United States, Mexico, much of Africa and China, and the Middle East. Other inputs to agriculture, especially topsoils, require examination everywhere, in a context of revised natural resource accounting (Repetto et al. 1987). MSUs and MSAs that pose the greatest constraints will determine the carrying capacities of any region in the absence of imports. Especially careful consideration must be given to assumptions about maintaining access to limiting resources through trade, because the last frontiers for acquiring cheap and plentiful resources are closing (e.g., Folke et al. 1991).

Because further degradation of the global environment is inevitable, interdisciplinary evaluations of the relative costs of alternative evils and their communication to the public is necessary. Some provision of insurance should be taken in proportion to the level of uncertainty and the severity of possible deleterious effects of given activities. In the meantime, no further net loss of essential elements of natural capital should be incurred.

Several potentially effective social (especially market) mechanisms have been suggested to make short-term incentives consistent with long-term sustainability. These mechanisms include fees for use of common-property resources, taxes on the depletion of natural capital, and flexible environmental assurance bonding systems for regulating activity that may be environmentally damaging, but whose effects are uncertain (Costanza 1987,
Costanza and Daly 1992, Costanza and Perring 1990). Implementation and further development of such methods of avoiding social traps is essential.

Frequently lacking, however, is a vision of a desired world that would establish a basic social carrying capacity for human beings. In the short run, efforts must be made to minimize the damage to Earth's systems, while providing the requisites of a decent life to the entire global population. In the long run, however, public discussions should be encouraged to guide policy on sustainable resource management. Sound science is central to the estimation of carrying capacities and the development and evaluation of technologies, but it can give minimal guidance at best regarding the issues surrounding the question of the kinds of lives people would choose to live.

The current decade is crucial, marking a window of environmental and political opportunity that may soon close. Environmentally, each moment of inaction further entrains irreversible trends, such as the global extinction of biodiversity and alteration of the gaseous composition of the atmosphere. Though it is certainly possible that intensifying human impact on the planet will precipitate a sudden disaster, it seems more likely that humanity will just gradually erode Earth's life-support capabilities over the next few decades. The more important window may thus be a political one for laying the institutional foundations for desired change. Right now, in the wake of United Nations Conference on Environment and Development, citizens and national governments may be at a peak in receptivity to acknowledging environmental problems and tackling their solutions. Let us seize the day.

Acknowledgments

We wish to thank S. Alexander, A. Ehrlich, M. Feldman, N. Haddad, H. Mooney, J. Roughgarden, T. Sisk, and P. Vitousek (Department of Biology, Stanford); L. Daniel and T. Daniel (Bureau of Economics, Federal Trade Commission); W. Falcon and R. Naylor (Institute for International Studies, Stanford); L. Goulder (Department of Economics, Stanford); J. Harte, J. Holden, and A. Kinzig (Energy and Resources Group, University of California, Berkeley); D. Murphy, A. Launer, and J. McLaughlin (Center for Conservation Biology, Stanford); P. Matson (NASA, Ames); S. Schneider (National Center for Atmospheric Research); and V. Valdivia (Siemens A.G.) for helpful comments on the manuscript. Holden, Matson, and Vitousek have been closely involved with us in the Stanford Carrying Capacity Project. Reviewers for *BioScience*, including R. Costanza, H. Daly, J. Miller, N. Myers, and one anonymous referee, were also extremely helpful. This project was made possible by the generous support of the W. Alton Jones Foundation and several private donors.

References cited


———. 1991b. Testimony before the Committee on Science, Space, and Technology (Hearings on Technologies and Strategies for Addressing Global Climate Change); US House of Representatives, 17 July.


770

BioScience Vol. 42 No. 10