

Chapter 2

Development of Environment Statistics in DMCs: Issues and Problems

This chapter reviews some common issues and problems that are likely to occur during the process of developing environment statistics in DMCs. It starts by giving a generalized picture of the different forms of environment data so as to facilitate a description of these forms in their common setting. The objective of the whole exercise is to delineate the role and possible contents of environment statistics as well as provide some insight into the derived uses of environment data.

Types and Uses of Environmental Data and Statistics

The major users of environment data and statistics are policy makers, planners, scientists, students, and the general public.

State-of-the-Environment Report

The position of environment statistics might be better appreciated by contrasting it with state-of-environment reports (SOERs). SOERs are not neutral instruments as they are explicitly meant to assess the condition of the environment, and determine the causes of the condition and possible cures. The text of an SOER gives an interpretation of existing data and knowledge so as to suggest what should be main trends. An example is ESCAP's SOE Report (ESCAP 1992). UNCED country reports can also be regarded as such (UNCED 1992).

Reporting on the state of the environment is becoming very common. Two types of products are usually produced by governments: SOERs and environmental statistical publications. The main distinc-

tion between the two is that environmental statistical publications are largely numerical, whereas SOERs contain substantial explanatory text. Scientists' need for detailed environmental statistics is obvious. But what is the purpose of the more general SOERs?

Policy makers need basic references on the state of the environment. Because SOERs are abstracted from a huge mass of reports and data, they do not normally contain all the environmental information available for a particular geographical region. However, they do provide a general picture of the state of the environment, from which progress in dealing with environmental issues can be deduced. In some countries, such as Italy, simply having to produce such a report has galvanized governments into action. SOERs also allow the environmental record of governments to be open to scrutiny. In developing countries, SOERs often represent the main source of information on the environment. Presently, many countries have environmental administrations, but not all publish regular SOERs. Of course, producing an SOER does not necessarily indicate that prudent environmental policies are being followed.

SOERs indicate where urgent problems may lie or where "unfinished business" remains. SOERs can also help policy makers to identify opportunities for improving the environment and to set priorities.

In essence, the SOER divides the big environmental problems into coherent specific problems and links them with significant actors (called target groups). Themes have been formulated for specific environmental problems such as the greenhouse effect, ozone depletion (global level), acidification (continental level), eutrophication (regional level), noise (local level), and target groups have been identified: households, energy plants, industry, agriculture, and car owners, groups to which more or less uniform measures can be applied. Further policies might be applied to specific areas, e.g., mountainous areas prone to erosion. Needless to say, an SOER for such an environment tends to lead toward a more specific and precise assessment of the goals laid out in the EPP.

Forms of Environment Statistics

The primary audience of environment statistics (and statistics in general, for that matter) is the policy makers. For the needs of this

audience, basic, straightforward environmental data are presented in derived forms. To clarify the importance of environment statistics, the most important forms (statistics, indicators, and indices), from basic to derived, for presenting environmental data are discussed. In practice, the forms are not mutually exclusive.

Environment statistics, briefly defined, refers to statistics on the state of the environment, as influenced by man's doings and natural causes, thereby showing causes, consequences, and remedial measures.

Such a definition is as broad as it is vague. In practice, environment statistics are a system of selected data organized on the basis of some framework so as to reflect certain natural or logical connections. Thus, all published environment statistics indicate some explicit or implicit framework due to the assumed connection and selection of items. The UN Framework for the Development of Environment Statistics (FDES) sketches the generalized traits of country publications (UN 1984). A unique feature of environment statistics is that they combine data from the natural sphere with those from the social and economic realms. For example, air pollution is presented as a chemical and physical phenomenon that is caused by man's economic activities and has impacts on man, plants, and animals alike.

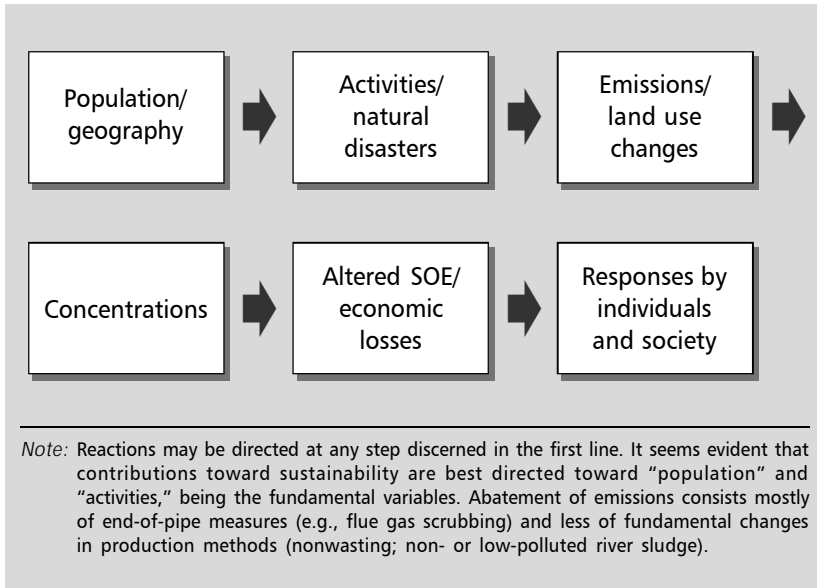
A framework should be seen as a tool for organizing and interpreting data statistics and, therefore, a framework cannot be a neutral device. Although a carefully developed framework is very important and useful for organizing and highlighting a country's special needs, it should be borne in mind that there is no ideal framework.³ A very general framework is reproduced in Figure 2.1. The framework basically describes the cause-effect chain.

Environmental indicators

Environmental indicators are an efficient way of measuring the environment issues in a country. Properly derived indicators can serve to highlight changes in environmental conditions that warrant

³ In contrast to economic statistics, which are guided by the system of national accounts and whose framework is based on a set of rigorously tested economic principles.

Figure 2.1
A General Framework for Environment Statistics



further investigations. Potentially, indicators can signal the health of the environment and can help in formulating actions to serve the long-term needs of the environment and the community. Environmental indicators can be classified into three groups: pressure, state, and response indicators. Pressure indicators show the causes of environmental problems. Certain flow quantities such as emissions, use of raw materials, products, and energy, or interventions in the environment, for instance, infrastructural activities that place a burden on the environment, are charted by means of pressure or stress indicators. State indicators reflect the quality of the environment in relation to the effects of human action. Response indicators pertain to measures taken by society to improve the environment. These indicators can tell whether things are getting better or worse, whether problems are growing, or whether current policies are achieving the desired goals.

It is not always possible to define an ideal state or norm for each indicator. Nonetheless, the indicator still provides useful insights into the trend of the state of an environmental resource and associated

concerns, if any. For example, it is difficult to establish how much ambient concentration of carbon dioxide (greenhouse gas) is acceptable. However, the increase or decrease in ambient carbon dioxide concentration gives an indication about global warming. By restricting the definition of an indicator to a simple structure as discussed, information provided by indicators would be objective. However, this information would be relevant to only one aspect of an environmental resource. Fragmented information provided by indicators may help in making environmental decisions, but does not provide a comprehensive view of the various dimensions of the environment for decision making. Environment indices attempt to overcome this particular limitation of environmental indicators.

Environmental indicators are not easy to formulate, and the amount of work involved in developing an agreed-upon set of indicators for a country should not be underestimated. The United Nations Commission on Environment and Sustainable Development (UNCESD), together with the UNSD, has already achieved some progress in identifying key indicators on the environment and sustainable development for the benefit of the member countries. Therefore, it is advisable that developing countries instead of reinventing the wheel should start developing indicators specific to their countries, proceeding from the work already done at the international level.

Environmental indices

An index combines a number of variables into a single value. The ability of an index to provide information at a level that encompasses information on a number of variables in the form of a single value makes the concept of an index attractive for a number of functions. An environmental index is necessary to reflect the state of an environmental resource; to understand the dynamics of an environmental system or the relationship between different environmental components as part of scientific investigation; to facilitate the analysis of trade-offs between objectives, i.e., development and environmental protection; and to assist in making resource allocation and policy decisions.

The components of an environmental index vary with the purpose or use of the index. If the task is a simple ranking of countries,

fewer problems are encountered than if the index is to be used for detailed comparisons. On the other hand, if one wants to compare the level of one country's environment quality with another's, then accurate and appropriate measurement techniques are critical. For an index to capture the environmental quality of a country, methodologies that take into account the discrepancies caused by differences in population, area, and income, among other factors, must be developed. Most important, however, is the selection of the set of environmental quality indicators that will be used in the index. Although the parameters that are chosen should be good indicators of environmental quality, it is most important that the availability of data be taken into consideration.

The Asian Development Bank has been promoting policies and programs to enhance the environment performance of its DMCs. The Bank's recent policies for reflecting the growing emphasis on environmental and social considerations in its funded technical assistance have brought about the need for new methods in assessing the environmental repercussions and the sustainability of economic development. The Bank therefore initiated a study in 1994 to develop tools for monitoring environmental change in the DMCs. A set of environmental indices, namely, cost of remediation, environmental elasticity, and the environmental diamond, were developed and tested in six selected countries of the region. All three indices were developed on the basis of four principal environmental components, i.e., air, water, land, and the ecosystem, using principal component analysis techniques. Each of the three new indices is designed to characterize different aspects of environmental quality. Each has different implications for data requirements. Without good-quality environment data, it would not be possible to make any judgment on the usefulness of these indices. In addition, the technical complexities and methods for constructing the indices also vary. The ultimate test for these indices is, however, their usefulness as tools for conducting future environment planning and devising effective policies.

Environmental accounting

The environmental degradation associated with economic development and population growth is visible in many countries of the Asian and Pacific region. The change in quality of land, air, and

water as well as the loss of flora and fauna makes one concerned about the costs of progress. Questions about the desirability of untrammelled economic development initially elicited the response "poverty is the worst pollutant." While there is much truth in this observation, our understanding of environment issues has become somewhat more sophisticated over the past few years. The environmental outlook of the developing countries of the region would be gloomy indeed had countries not been alerted now to the need to protect their environment and natural resources. While the significance and the degree of severity of these environmental problems vary from country to country, their mitigation inevitably requires the integration of environmental considerations into the planning process and developmental activities.

The preparation of environmental accounts and their regular publication could bring a much greater degree of accountability in public policy. Environmental accounting aims at measuring the real income of a nation, which takes into account how much nations borrow (or take) from nature. Changes in environmental quality and stocks of natural resources that occur as a result of economic and social development must be taken into consideration so that development decisions can satisfy the needs of present and future generations.

The ongoing exploration of the usefulness of accounting techniques for the organization and compilation of aggregate environment statistics gives rise to a multitude of different approaches. These can be classified into either physical or monetary accounting schemes, but many countries in the region appear to be faced with the need to advance in both types of accounting frameworks. It seems that the demand for monetary accounting primarily emanates from political quarters and some economists and environmentalists, while the more technically oriented subject-matter experts appear to be sceptical of the possibilities for aggregation, for monetary valuation, or for both. Demands for the compilation of accounting aggregates are more often than not associated with quests for developing statistics permitting the conceptualization of sustainable development.

The conventional system of national accounting does not adequately reflect the effects of a degrading environment on the economy, or of depleting stocks of natural resources; indeed it is not designed to. Various approaches to environmental and resource accounting have been proposed to deal with the shortcomings

identified in national accounting systems with respect to the environment. The revised 1993 System of National Accounts (SNA) considers environment as a “satellite account,” separate from the core accounts, but developed in a framework entirely compatible with the SNA. The UNSD has prepared the *Handbook of Integrated Environmental and Economic Accounting* (United Nations 1993), which provides guidelines for producing “satellite” integrated environmental accounts. The handbook describes a system of integrated environmental and economic accounting (SEEA) that was tested and revised following case studies in Mexico, Papua New Guinea, and Thailand. Essentially, the new satellite system provides guidance in the treatment of natural capital depletion (land, water, air). For goods and services that have no market prices, alternative valuation methods are suggested to impute the cost of depletion or degradation. Using SEEA, planners incorporate environmental information into production, income, and balance sheet accounts. It explicitly allocates all environmental impacts to the separate economic activities that cause (or bear) them. Environmental information is thus made available for integration into the entire array of economic policy and management analyses that the national accounting system serves.

The UN satellite environmental accounts will certainly encourage interested countries to begin constructing environmental accounts that can easily be integrated into the core system, but they are yet to be fully tested. The few case studies undertaken were based on far-from-perfect data — sometimes only rough estimates, though reasonably comprehensive — focusing on minerals, forests, and some pollution charges. As the countries where the methodology is tested come to some agreement about what the international standard of green accounting should be, the possibility of reforming the core UN system to include environmental information will grow.

Contents of Environment Statistics

UN-FDES wisely abstains from presenting model contents for environment statistics, but restricts itself to presenting a general framework within which actual concerns are listed. But for one faced with the problem of devising a handbook, breaking up the field into definite chapters is inevitable. Furthermore, as people are unfailingly

bewildered about the possible scope of environment statistics, a broad sketch of the contents of environment statistics may be helpful. It should be stressed that the following is not in any way meant to be a prescription, but is merely a tool to provide clarity on the subject.

Core statistics

Core statistics, which are considered the main subject matter of environment statistics, comprise the physical statistics describing emissions from human society into nature as well as changes in land use and their consequences. They cover the following areas of concern:

- (i) emissions to, concentrations in, or effects on air, water, soil
- (ii) land use, soil degradation, deforestation, agriculture, or salinization
- (iii) natural environment: plants, animals, or ecosystems

At present, the best developed methodology is for emissions and concentrations, the latter often referred to as quality data.⁴

Probably for historical and institutional reasons, emission and concentration monitoring developed independently. Thus the opportunity to link the two and thereby provide policy makers with more coherent data was lost. For example, one may know that a river is polluted, but if one cannot identify and quantify the relevant pollution sources, corrective measures are bound to be vague and probably even misguided.

At a later stage, when sufficient data are available, data for different media (air, water, soil) could be integrated by following the fate of specific substances in so-called substance balance sheets. Many effects can be readily measured, but their relation to causes can often be stated in general terms only. One example is the incidence of lung diseases and air pollution.

⁴ Between emissions and ambient concentrations, dispersion and emission can be interposed. Dispersion is the way emitted substances are distributed from their point of release. Emission is the process in which substances pass over to another medium where they may reside for some time, e.g., air emissions to soil, water emissions to soils.

As most developing countries are still highly dependent on agriculture, the importance of land use or land cover statistics in connection with soil degradation and other soil-related problems can hardly be underrated. In countries poor in minerals, oil, gas, and the like, soil is the most precious resource. Unfortunately, no unified methodology has been developed for assembling reliable data in a common framework. However, efforts being exerted by FAO and others appear promising (Earthwatch/GEMS 1994).

Statistics on the natural environment are strongly related to land use statistics (in this case, formulated in terms of habitat and ecosystems) as a consequence of changing land use patterns rather than of pollution. And much like effect statistics, many data can be produced (e.g., abundance and distribution of plant and animal species) without their making much sense for lack of an appropriate framework, even on a national or subnational scale. Meanwhile, species inventories and reed lists seem to have less meaning in the tropics than in the higher latitudes, as there are simply too many species to count. Ecosystem protection might be a better option in this case. The most important questions then are (i) Do current protected areas match the threatened ecosystem? (ii) Have boundaries of protected areas been suitably delineated? (iii) Are enough areas under protection? (Braatz 1992).

Explanatory and background statistics

Starting from the assumption that environmental problems are caused by man and his actions or activities (Figure 2.1) and modified by natural circumstances and wealth (amount of production and consumption), the following statistics should provide explanatory variables to emissions, concentration, or land use statistics:

- (i) population (absolute level, density, rate of change);
- (ii) activities (including transport, energy, and elements of SNA); and
- (iii) geography and climate (including natural disasters).

The methodology for these statistics has already been established (e.g., population statistics, SNA). The statistical processes consist mainly of selecting and restructuring existing data, in the light of some framework for environment statistics. The same applies

to geographic and climate data. For natural disasters, for example, WMO methodology and data may be used for a start.

Integrated statistics

There is concern that categorizing environmental information by subject areas such as air, water, land, etc., may lead away from an integrated approach to environmental issues. It is necessary to recognize that the environment and the economy are closely interrelated. The databases that are developed should be based on the dynamic real world and not on an arbitrary set of subject areas.

Statistics on environmental costs and physical resource accounting are most useful in a direct way to developing countries. The first consists of surveying the outlays made to prevent, or restrict, environmental pollution and damages and to restore the environment. Physical resources accounts represent gains and losses in the resource stocks (oil, iron ore, wood, fish) in the course of a year, and thus offer additional information to the SNA.

Statistics in environment frameworks

Judging from the work done by UNSD, UNEP, and Habitat, the following frameworks seem to be important to developing countries:

- (i) human settlement statistics linking socioeconomic data and statistics with environmental services and infrastructure; and
- (ii) urban area statistics.

In contrast to integrated statistics, such statistics offer a cross section of several statistics in a more or less defined framework. At the lowest level, it may consist of a national ordering of existing statistics. The combination of social variables with environmental ones is typical, while human settlement statistics seem to stress linking the social with environmental variables. At the household level (family composition, income, education, health, versus access to water, cooking fuels used, waste disposal, etc.), urban statistics tend to concentrate on environmental problems typical to densely populated

areas like megacities, including spatial planning issues. Noise, traffic emissions, and waste generation and treatment pose a much greater problem in cities than in rural areas. Furthermore, more and better environmental data may be expected to be available for cities than for rural areas (ESCAP 1993).

Framework for the Development of Environment Statistics

As discussed earlier, environment statistics are multidisciplinary in character. Their sources are dispersed and a variety of methods are applied in their compilation. Better coordination and organization are needed for developing this complex area of statistics. To this end, statistical frameworks and systems have been successfully applied in the fields of social, demographic, and economic statistics. A similar framework would be useful to provide a systematic approach to the development of environment statistics.

The statistical framework is an instrument by which data coming from various data-collecting institutions are compiled and integrated in such a way as to be more useful in the formulation and valuation of socioeconomic and environmental programs and policies. Different conceptual frameworks are in vogue in different countries. A preliminary review of all those environment statistics frameworks suggests that four conceptual models are being used:

- (i) environmental media-based framework,
- (ii) resource accounting model,
- (iii) ecological model, and
- (iv) stress-response model.

The media-based framework organizes environmental issues from the perspective of the major environmental components of air, water, land or soil, and the man-made environment. It aims at assessing the state of environmental media at different points in time rather than at monitoring the processes of environmental change continuously. The media approach complies with conventional statistical and administrative concepts and classifications, and the popular perception of the environment. The main criticism against the media-based model is that it is more concerned with the natural

environment than with the human aspect, and that the interrelationships among the components are not considered.

The resource accounting model keeps track of stocks and flows of natural resources from the time of their extraction from the environment, through successive stages of processing and final use, to their return to the environment and waste or to the economic sector for recycling. This system has also been made compatible with the 1993 SNA. Although this model appears to be appealing, theoretically there is difficulty in implementing it to trace the complete life cycle of resources, a task that requires a high level of coordination among various agencies.

The ecological approach to statistical data collection and analysis, on the other hand, includes a variety of models, monitoring techniques, and ecological indices. They deal with such diverse topics as the assessment of population diversity and dynamics, of biomass production; and of the productivity, stability, and resilience of ecosystems.

The stress-response model was developed in recognition of the inadequacy of the media approach for describing the processes of environmental change. The stress-response model focuses on impacts of human intervention within the environment (stress) and the environment's subsequent transformation (environmental response). By establishing cause-effect relationships between activities and subsequent environmental impacts, the framework helps in developing statistics that are useful for taking both preventive and curative measures for protecting the environment and mitigating the adverse environmental impacts of development activities. Many existing frameworks are broadly based on this conceptual approach. The main disadvantage of this approach is that it represents a simplistic view of the environment by ignoring interactions among environmental components. The nature and severity of the impact of an activity on an environmental resource could depend on both the existing status of that resource and the interactions that this resource has with other environmental components. In fact, the interactions are quite complex, and, as a result, one impact could trigger a chain of other environmental impacts.

The United Nations Framework for the Development of Environment Statistics (UN-FDES) is based on the pressure-state-response model, which attempts to establish a relationship between

human activities and the resulting environmental impacts. This approach guides the (i) development of a statistical system for measuring human activities that cause stresses on the environment, and (ii) measurement of actual impacts that have become distinct over time and space. In addition, it takes the policy response, mitigating measures, and stocks or inventories into consideration.

The UN-FDES was designed with a view to assist countries in the development, coordination, and organization of environment statistics. The use of such a framework is envisaged for these specific purposes:

- (i) review of environmental problems and concerns and determination of their quantifiable aspects;
- (ii) determination of variables for statistical descriptions of the quantifiable aspect of environmental concerns;
- (iii) assessment of data requirements, sources and availability; and
- (iv) structuring of databases, information systems, and statistical publications.

The FDES format has evolved from the joint consideration of the scope and nature of environment statistics, and the purposes and properties of such a framework. A synthesis of these factors has resulted in the format of a two-way table that relates the basic components of the environment to the various categories of information (Figure 2.2).

The components of the environment are based on the description of the coverage of environment statistics as perceived by developing countries. The natural environment includes the environmental media of air, water, and land or soil, as well as the biota in these media. The man-made environment includes human settlements, which consist of the physical elements shelter and infrastructure, and services to which these elements provide the material support. The information categories are anchored on the recognition that environmental problems are the result of human activities and natural events. Relevant information refers to social and economic activities and natural events, their impacts on the environment, the responses to these impacts by governments, nongovernment organizations, enterprises, and individuals.

Figure 2.2
Format of the FDES Framework

Components of the Environment	Social and Economic Activities, Natural Events	Environmental Impacts of Activities/ Events	Responses to Environmental Impacts	Inventories
<ol style="list-style-type: none"> 1. Flora 2. Fauna 3. Atmosphere 4. Water 5. Land/soil 6. Human settlements 				

The usefulness of any FDES could be determined by the following factors:

- (i) adaptability of the framework to the existing institutional frameworks and mechanisms of the countries;
- (ii) organization of information to facilitate the flow of data from data-producing agencies to the user agencies;
- (iii) provision in the framework to preserve the interlinkages among statistics compiled for different environmental domains; and
- (iv) provision in the framework of a logical framework for identifying relevant agencies that will compile nontraditional statistics, the need for which has arisen because the interactions among various environmental components have become important.

Among the various frameworks discussed here, the UN-FDES, despite its limitations, has been found most operational. This framework has already been tested in some countries, and the results are satisfactory. One of the objectives of the Bank's RETA is to establish a sound foundation for an environment statistics system, so that environment statistics are compiled on a continuing basis by the

participating countries. During the inception meeting of the RETA, the selection and adoption of an environment statistics system were discussed in detail by the participants. In view of the framework's simplicity, flexibility, and adaptability, the RETA participating countries decided to pattern their national framework on environment statistics on the UN-FDES. It does not mean, however, that the countries are required to adapt this framework forever. The frameworks will need to be modified over time to reflect the increasing and changing demands of environment statistics.

Compendium of Environment Statistics

When the national framework for the development of environment statistics is finished, the next logical step for any developing country will be to prepare a compendium of environment statistics based on that framework. A large amount of environment-related statistics are collected in the countries on an ad hoc basis. These statistics need to be organized systematically to make them useful for policy makers and decision makers. The compendium should attempt to identify statistical parameters and variables for each component and information category of the environment as laid out in the national FDES. The objective of the compendium should therefore be to (i) improve the utilization and value of existing information of the NSO and other environment agencies; (ii) provide a comprehensive, easy-to-understand reference of environmental information for the country; (iii) provide a directory of environmental information sources; (iv) raise awareness and understanding of the environmental issues of the broader community including government, industry, nongovernment organizations (NGOs), media, and the wider public; (v) set up a process of cooperation, consultation, and coordination among data providers of the country for the ongoing development of environmental information in an effective and cost-efficient way; and (vi) develop the skills of the NSO staff in compiling, collating, and presenting environmental information.

Collecting data on each of the areas of environment could be cumbersome and resource demanding. The best action will be to start compiling data from existing sources. Different environment-related agencies undertake a large number of surveys or studies, often without

much coordination among themselves; thus, they contribute to the duplication of activities in many developing countries. These agencies also compile a large amount of environment statistics for implementing various environment-related projects. It is clear that a large amount of environment-related statistics could be compiled by scanning various ad hoc studies and administrative records of the government agencies.

The presentation of environment statistics is not among the most important considerations when starting; what should be stressed is the use of the explanatory text to avoid misinterpretation, to introduce technical matters, and to point out links with other statistics, among others.

The FDES prepared by the participating countries has identified a long and comprehensive list of variables for each component of environment and information categories. They have also identified units of measurement, computation levels, data availability, sources of data, and data collection methods for each of the identified variables. The outcome will provide a comprehensive knowledge of the status of environment data in a particular country. What is important now is the organization of future studies and surveys in a manner that the collection of the important environment-related information becomes systematic and data can be updated on a regular basis. It would also seem important for each NSO to examine the possibility of including some relevant environment-related questions in the schedules of various surveys that the office undertakes on a regular basis.

State-of-Environment Statistics in Developing Member Countries

Robust environmental and socioeconomic data provide the foundation for the analysis and interpretation of the state of the environment. In the absence of such data, any report on the state of the environment is reduced to a descriptive, anecdotal, and nonsystematic observation, which is not an acceptable basis for rational decision making. The type of data required covers a wide spectrum. Data on natural resource stocks and environmental conditions are essential. Similarly, statistics on human activities impacting on the environment, the

emission of pollutants, natural events, and human responses to environmental changes are equally important in assessing the ecosystem interactions.

Environmental and socioeconomic data tend to be collected independently by diverse agencies, using different methods and classifications, and for quite specific purposes. Data on the environment are usually derived from monitoring programs and the interpretation of remotely sensed images. Socioeconomic data are collected from statistically designed surveys and from administrative records. From a state-of-the-environment perspective, particularly at national, regional, and global levels, the spatial resolution and temporal dimensions of much of the data are often limited. Much of the available data relate to individual environmental or human activity components rather than to a synergistic, ecosystem perspective. For example, databases for commercial forest areas tend to emphasize the production mandate of the forestry management agency, not adequately reflecting the diverse values of forest ecozones, which include their role in terms of habitat and biodiversity, water conservation, and traditional and alternative land uses.

Under these circumstances, do the countries need more data or should they use existing data more efficiently? Some people would contend that the developing countries have insufficient data, others think that they have too much. We know comparatively little about the biota of the planet, particularly microorganisms; yet masses of raw data come to earth daily from observation satellites. Our databases for social, economic, and demographic conditions and trends are relatively well developed and integrated compared with that for environmental data. However, it is clear that we do not have adequate and accurate data to answer some basic questions related to sustainability. Data limitations in terms of both balance and quality severely hinder the quantitative assessment of and reporting on the state of the environment. In today's world, data must support the development of more holistic information, understanding, and knowledge (Rump 1996).

Although high-quality data are vital for credible information, a systematic approach to their generation is largely lacking. The acquisition, processing, and storage of environmental data is time-consuming and expensive, and is not a priority for most governments. Consequently, baseline and trend data related to the ways the

ecosystem functions and its components interact are insufficient. The data we do have tend to be scattered and difficult to obtain, while proprietary and security factors can inhibit dissemination and open access. Environmental and socioeconomic data do not generally exist in usable and integrated formats for reporting. There is a common deficiency of infrastructure and standards to facilitate the easy exchange and correlation of data from different jurisdictions and disciplines.

What can the environmental agencies do to improve the situation, recognizing that they are usually not agencies for gathering primary data? The various sources of environmental data include bodies responsible for environmental and resource monitoring, statistical surveys, mapping, and remote sensing. There are no quick solutions, but many agencies are taking immediate and longer term steps to consolidate data to facilitate more efficient monitoring and reporting. There is a basic choice between (i) designing from the outset a new monitoring and data processing system based on ecosystems, to allow the reporting of environmental conditions and trends; and (ii) incrementally adapting existing monitoring and survey systems to meet today's more holistic needs. Both approaches have been proposed. The suggestion in this chapter tends to be modest and is primarily directed at building on existing data and data-gathering networks to improve access and dissemination. This approach is seen as being more practical and more suited to the resource realities faced by most national and regional reporting programs.

The collection and compilation of environment statistics constitute a recent phenomenon in most of the developing Asian and Pacific countries. The present system of data collection in all these countries is weak, unorganized, and poorly funded. Most of the countries do not collect core environment statistics; what they have is environment-related statistics. Where some core environment data exist, their quality, comparability, and accessibility normally fall short of the standard required for decision making. There is a wide variance among countries with respect to the extent of their expertise and knowledge. There are also variations in their interpretations of terminologies, classifications and standards, estimation methods, the training they provide to their personnel, and the resources they allocate for data collection.

Despite some data gaps, there have been some efforts by international organizations to compile and publish global/regional environmental data. Since there is a general dearth of environmental data in all the developing countries, it is always likely that international data will show gaps. For example, no international organization reports the extent of tropical deforestation on an annual basis, primarily because most countries do not make annual assessments of their forest resources. FAO assesses deforestation only once every ten years. No amount of international efforts can therefore succeed in compiling regional/global statistics unless the countries' capabilities to produce environment statistics are improved.

Problems and Issues in the Collection of Environment Statistics

The organization of environment statistics offers challenges to the statisticians who are highly familiar with the techniques of collecting socioeconomic and demographic data, but not so much with the techniques of collecting environment data. Since the environment is all-encompassing and ill understood, an ongoing debate among statisticians is whether they should be publishing imperfect or incomplete data. Most national jurisdictions and international organizations, however, define environment data to include a broad spectrum of information, thereby implying that, excepting certain autonomous natural events, the environmental changes of concern to citizens and governments are the result of human activities.

Some Methodological Issues

This section aims to give readers an overview of the scientific issues and complexities that arise during the measurement, interpretation, and presentation of environmental data. A more detailed description of these issues is given in Chapter 4.

Measurement errors

There are certain important differences between traditional socioeconomic data and physical data about the environment. In

socioeconomic data, most of which are obtained from surveys, the measurement error relates to questionnaire design and the understanding of the human cognitive processes. Quantifying the differences between what one intends to measure and what one really measures with a survey may be a problem. On the other hand, most environmental data are measured using scientific instruments for which calibration standards can ensure repeatability.

Many air quality statistics are usually obtained through monitoring. This implies that statistical generalizations from the data are impeded by the site-specificity of monitoring data, the characteristics of measurement sites, sampling design and specification, calibration, and sampling and analytical methods used. The repercussions of these factors on data reliability and comparability are known only to some extent. A remedy for this situation could be the incorporation of statistical concerns in the design of monitoring systems, such as the selection of monitoring stations with a view to obtaining spatially representative measurements. While some progress along this road is slowly materializing in some countries, the overall situation remains unsatisfactory. Future works should bring about clarification on this respect. Of special interest are studies into the relationship between the site characteristics and levels of measured data. Furthermore, future methodological studies should shed further light on the possibilities for valid generalizations of monitoring data to air quality statistics (Hamilton 1991).

Sampling

The question of how to sample from the universe of industrial establishments is a particular methodological concern, and a number of approaches have tentatively been used. Large industrial enterprises are primarily responsible for the generation of the bulk of wastes; they ought to be adequately represented in the samples to permit the making of valid estimations. On the other hand, particular types of waste may only be generated in small or medium-size forms so that samples confined to large units would disregard the types of waste concerned. A relationship exists in this regard between the details asked in the statistical inquiry and the importance of the question. It appears that the use of fairly detailed lists of wastes in a survey calls for the inclusion of small and medium-size

establishments in the sample. It is hoped that further insight into the question of how to sample will become available once relevant work in progress in several countries is carefully evaluated in this regard.

In collecting socioeconomic data, simple statistical techniques can be used to compute relevant basic statistics. These industrial systems allow the use of the enumerative measurement techniques for collecting data. In principle, measurements for such systems can cover the entire population, as these systems are made of distinguishable statistical units. Sample surveys are conducted at regular intervals to measure other socioeconomic parameters, such as the economically active population, household income and expenditure, and employment. These surveys are based on probability samples, with sample size ranging from hundreds to tens of thousands of households. Large sample sizes for these surveys allow the use of the Central Limit Theorem to make the assumption that sample averages follow normal distribution. Therefore statistics computed from these samples could be used to make an inference about the entire population within a known error bound.

However, descriptive statistics cannot be used for environmental resources such as the atmosphere, water, and ecological systems. Since air and water resources are contiguous and measurement units are not distinguishable, samples of air and water are to be taken for the measurement of pollutant concentrations in these resources or media. Statistics computed from these samples could be used to make an inference about the air quality of a region or the water quality of a water resource. For ecological systems, samples also need to be taken for identifying and counting main taxa, to make an inference about the population of flora and fauna in a region. Measuring the entire population of the region would be prohibitively expensive (ADB 1995).

For ecological systems, it is even more important to take into account spatial structure and dynamics when selecting the sampling locations or sites. Ecological systems are complex systems made up of an assemblage of communities and abiotic elements, continuously interacting with each other. Hence, statistics on these systems deal with both structural and functional measures. The selection of sampling locations should capture information on community frontiers and allow monitoring of the movement of these frontiers. Sampling

sites should also take into account the changes close to the boundary of the ecosystem to monitor the impact of human settlement pressures or the influence of the buffer zone on the expansion of the ecosystem.

Ignoring the spatial structure of ecological systems when selecting random sample sites will result in the loss of information on the structure of an ecosystem. Surveys have shown that the ecosystem may be divided into several zones to carry out stratified random sampling. For developing reliable statistics within each stratified zone, a probability distribution function of the parameter or variable of interest should be known to determine the minimum number of samples required to provide statistics of a given confidence level. In principle, it may be possible to determine the probability distribution function and the minimum number of samples for computing reasonably reliable estimates. However, budgetary constraints may limit the number of samples and thereby influence the reliability of the statistics.

Aggregation

Many socioeconomic data are extensive (number of individuals, value of shipments, etc.), and so may be summed into aggregates. A considerable proportion of environmental data are intensive (concentrations of pollutants in water, air, or soil, etc.) and therefore may not be summarized or averaged without auxiliary information. While environmental change and its causes may be global, a significant proportion of environmental problems are localized in space or time. Therefore, in most instances, each data in an environmental database should have a geographic reference. For reporting environmental data, natural geographic boundaries such as ecological zones, watersheds, soil and climatological regions, wetland boundaries, forest zones, etc. may be preferable to administrative zones. However, introducing geographic location as an attribute of data may open up whole new areas of potential statistical error, including the accuracy of geographic referencing and of summarization of data to environmental boundaries (Hamilton 1991).

Natural environmental resources also show spatial and temporal variations in their various characteristics. These variations

can arise due to both natural and human-induced forces. For example, uneven distribution of rainfall over a region can be attributed to natural forces. Spatial variations in air and water quality, on the other hand, can be due to the uneven distribution of pollution sources as well as natural factors like wind direction, speed, and topography for air quality; and the volumetric flow rate of a river and river geometry for water quality. The scale of temporal variations for the atmosphere can be as short as a few seconds or as long as a year. For a water resource, the scale of temporal variations is determined by the pattern of pollutant discharges into the water body as well as by seasonal and annual weather cycles. For ecological systems, the scale of temporal variations could be several decades, as is evident from ecological succession.

Both spatial and temporal variations need to be captured by statistics in such a manner as to provide useful information with respect to the concerns about the uses of these environmental resources. For example, to monitor the trend in global warming, both spatial and temporal aggregations of greenhouse gas concentrations are required. Spatial aggregation is required over a large area in such a manner that the statistics would be free from influences of peak concentrations due to individual sources. To capture all scales of temporal variations, aggregation must also be carried out for measurements taken over a year. On the other hand, for statistics representing exposure of the population to pollution, aggregation is required to determine the maximum one-hour, eight-hour, daily, and annual pollutant concentrations to reflect the risk from both acute and chronic health effects.

If spatial and temporal variations of resource attributes are of a random nature, or if these attributes follow normal distribution, random samples taken in space and time could be used to infer statistics on the environment resources, or a part thereof. However, the selection of random samples without considering the heterogeneity of the system due to natural or human-induced forces could result in loss of valuable information. For example, spatial aggregation of rainfall data based on random samples taken over a large area without considering the differences in climatological or rainfall regimes would provide statistics that are useful neither for agricultural planning nor for the management of water resources (ADB 1995).

Spatial aggregation

Spatial aggregation of data is required to compute statistics that are representative of an area domain, such as the sulfur oxide concentration in an urban area, the dissolved oxygen level of a lake, rainfall intensity in a watershed, or the species diversity of an ecosystem. To preserve the heterogeneity of the resource while computing statistics, areas within a region need to be delineated in such a manner that for the parameters or variables of interest, the variation within each area is low, while variations across areas are high. Samples taken from each area can then be used to compute statistics representative of that particular area. If the central tendency of a variable or an attribute over the entire region is required, statistics compiled for individual areas may be used to determine the median or the arithmetic mean, as may be appropriate for the type of distribution in the area.

Ideally, the delineation of heterogeneous areas should be based on data collected over a long period of time with the use of a dense monitoring network. These data can be analyzed using such techniques as spatial correlation analysis, cluster analysis, or principal component analysis to delineate areas that show significant differences with respect to given attributes of an environment system. For atmospheric systems, representative statistics for an area could also be determined by plotting contours of equal attribute values on a map representing a region. Thus, the average pollutant concentrations represented by two consecutive isopleths could be used to represent pollutant concentration in the area enclosed by these isopleths. Pollutant concentrations estimated in this manner could be used to determine the exposure of the population living in the area. Similarly, average rainfall intensity represented by two isopleths (lines joining equal rainfall points) could represent the rainfall of the area enveloped by these isopleths. The representative rainfall intensity could be used to calculate surface runoff over the area enclosed by isohyets.

Temporal aggregation

The temporal aggregation of data should capture temporal variations in the variables in such a way that the statistics can provide

information for different purposes. For the atmosphere, the scale of temporal variations ranges from minutes to years. With regard to air quality, minute-to-minute variations are caused by changes in wind direction and speed. Daily variations in air quality result from diurnal variations of the atmosphere as well as from daily emission patterns. Macroscale weather fluctuations last a few days and are important from the viewpoint of air pollution episodes. Seasonal and annual weather cycles also influence the air quality of a place. Temporal variations in the quality of a water resource may be caused by the daily pollution discharge pattern. The dilution capacity of a water resource may change in response to seasonal and annual weather cycles, which give rise to variations in water quality. Hence, the aggregations of air or water pollutant concentrations over different time scales may be required to derive statistics for different uses. For example, it may be necessary to determine episodic air pollution levels to shut down some industrial units in a region that is susceptible to air pollution episodes.

Similarly, it may be of interest to determine peak water pollution levels and the volumetric flow of a water body during the dry season for water resource planning. If pollution levels in air and water are continuously monitored over time, average concentrations over time intervals and their frequency distributions could easily be determined. However, cost constraints may prevent the continuous monitoring of all pollutants at all monitoring stations. Therefore, knowledge about the frequency distribution function is required to determine the frequency of sampling necessary to obtain reliable average concentrations for different time intervals. Budgetary constraints may not allow the monitoring to be carried out at the frequency determined on the basis of statistical analysis (ADB 1995).

Rapid Assessment Method

It is clear from the preceding discussion that there is a lack of core environmental data in most DMCs. In the face of this inadequacy or general lack of both quantitative and qualitative environmental data, undertaking environmental planning and management measures as well as formulating the requisite environment policy will be rather difficult. The treatment of water, air, and soil pollution cannot be compartmentalized. It is necessary

to approach these pollution problems from an integrative perspective. However, reliable data on these pollutive emissions and discharges in most cases are rather scarce because the requisite pollution monitoring for generating such data is both difficult and costly. Such monitoring assumes the existence of an organization with highly trained and skilled technical staff. For developing countries, therefore, the intensive use of these actual air quality measurements for the formulation of a preliminary emission inventory may not seem to be a valid option. Thus, field data for air quality monitoring may have to be collected only for restricted areas and selected air quality parameters and variables. In view of this, an indirect approach to air quality monitoring through emission inventories is therefore suggested. This approach, which is called the Rapid Assessment Method (RAM), was developed by the World Health Organization (WHO), especially for the benefit of developing countries.

RAM provides an effective way of assessing air, water, and solid waste emissions and discharges generated by each source, or groups of similar pollution sources, within the study area. It permits convenient assessment of the efficiency and effectiveness of alternative pollution control options. The method is based on documented and often extensive past experience with the nature and quantities of pollutants generated from each pollution source, with and without associated pollution control systems.

The main advantage of this method is convenience of use. It permits the conduct of integrated inventories of air, water, and land emission sources in highly complex situations within a reasonable period of time and at fairly modest resources. Despite the simplicity of the method, the end result has often been found relatively reliable compared with that of a direct source monitoring program, especially in cases where shortcuts to data collection have been taken. Another significant advantage is the possibility of estimating the effectiveness of alternative emission control schemes for their emission reduction potential. The latter constitutes a major input for formulating rational pollution control strategies (ESCAP 1998). However, a major disadvantage of the RAM approach is the statistical validity of its inventory predictions. More specifically, the predictions for any given pollution source will have to be considered as only indicative in many cases on account of the significant variation evident in normalized emissions between similar pollution sources.

Consequently, control measures adopted immediately following the application of RAM need to be viewed as preliminary, and subject to subsequent and more detailed analysis prior to implementation of pollution control strategies.

For any given industrial activity, the emission factors vary from one pollution source to another, the variations sometimes being significant. Such variations are often the result of different operations practices, but may also reflect differences in the design and layout of the process equipment. The emission factors provided are usually selected to represent average or typical conditions as much as possible. As a result, the calculated emission for any individual pollution source can be expected to occasionally differ significantly from the actual waste loads operated. However, overall pollution load predictions for a number of similar industrial plants, e.g., the total pollution loads in the effluents of many tanneries operating in a given area, should be reasonably accurate.

The question that often arises is on the validity of these emission factors across different countries, especially when those derived from industrialized countries are applied to developing countries. For example, because of differences in source inspection and the maintenance of specific industrial plants, or because of differences in the size of a "typical" industrial plant, the use of somewhat higher emission factors in developing countries could be justified. However, the extensive use of RAM for over a decade in many parts of the world has shown that this is not a significant problem (WHO 1992). The general conclusion so far has been that the application of RAM can be expected to produce acceptable accuracy for environmental planning and management purposes. Such accuracy could be improved in cases where information about local emission factors is available, with the requisite pollution assessments being derived from them whenever possible. Such refinements, along with the increase in the number and quality of trained and experienced personnel, are expected to improve results that could serve as useful inputs for purposes of environmental quality improvement and management.

RAM is a technical means to gain fairly quick insights into the various significant aspects of total pollution load, in terms of emissions and discharges, exerted by anthropogenic activities. The widespread use of this methodology is expected to lay the groundwork

for better environmental policies, and improved reporting of local, national, and international environment statistics. However, to make finer distinctions between data sets and to facilitate the formulation of time series data sets that are more closely reflective of real world emissions, the use of RAM should also be complemented by additional instruments that will ensure that emissions of large polluters (refineries, etc.) are monitored more directly; also, collective emissions by diffuse sources (i.e., from traffic) should be reflected more faithfully in environment statistics. What this implies therefore is that the RAM results themselves offer a good starting point for the undertaking and formulation of emission and discharge inventories that are adapted to the specific needs and circumstances of various countries.

Institutional Problems

Environment statistics cut across many sectors and subsectors. A number of agencies are involved in the collection and compilation of environment-related statistics. Lack of coordination among these agencies may hinder the collection and compilation of environment statistics. Hence, it is necessary to think about various strategies and mechanisms to achieve effective coordination between data-collecting agencies. National statistics agencies have traditionally been responsible for collecting, compiling, and disseminating socioeconomic and demographic statistics in all countries. Therefore, it would seem logical and most appropriate to entrust the responsibility for compiling and collating environment-related statistics as well to the NSOs. To provide guidance to the NSOs and to improve coordination among the data-collecting agencies, it is necessary to establish a high-power steering committee consisting of heads of these agencies. Such a committee will be helpful in sorting out any differences in sharing the data from the various agencies.

Environment being a highly specialized subject, statisticians are not quite well-versed with the concepts and technical terms of environmental science. To enable the NSOs to effectively collect and compile environment statistics, it is necessary that they acquire familiarity with the basic concepts and definitions in environmental science. Close interaction between statisticians and environment scientists is necessary. Creation of a number of technical committees will be helpful not only in providing technical guidance to the staff

of NSO, but also in maintaining working relations between the staff of NSO and other environment agencies. The collection of environment statistics will be facilitated if separate environment statistics units are established in the NSOs and are manned by well-trained professionals. These could be statisticians with adequate training in environmental science, or environment experts with some training in statistics. Since the field of environment statistics is new, regular in-country training in environment statistics should be organized utilizing the services of the various experts available in the countries. To meet the growing demand for trained manpower in this specialized area, such training could be supplemented by regional and international training programs, as and when necessary.

It is important to develop environment statistics within the context of regional collaboration. This can be achieved by the exchange of environment statistics reports and other products such as discussion papers and survey designs, and through the exchange of views and experiences at regional meetings and workshops. The work of the NSO of a country could benefit from the examples of other countries. The NSOs should establish contact with colleagues in the region for the purpose of acquiring copies of those countries' technical reports. The NSOs would also benefit from examples from countries with more developed environment statistics programs such as Australia, Canada, and Netherlands. Access to Internet may also help NSOs to keep abreast of international developments in the field of environment statistics.

There is general lack of adequate financial resources for the collection and dissemination of statistics including environment statistics in most countries. Thus there is a need to draw attention to this deficiency and to urge governments to take the necessary steps to improve the capability of the agencies involved in the collection, compilation, dissemination, and analysis of environment statistics.

Some countries of the region are more advanced than others in developing environment statistics. Therefore, the latecomers can really benefit from the experience of the early players of the game. There is, however, a tremendous communication gap among different countries of the region, among international organizations, and even among different divisions of the same organization on their respective activities in the field of environment statistics. It is desirable that

there should be reciprocity in information exchange among the concerned organizations so as to prevent duplication of efforts and ensure a more efficient utilization of resources.

Conclusion

The preceding discussion attempts to point out some of the methodological issues that may be confronted in the collection and compilation of environment statistics. The preparation of the national frameworks and compendiums of environment statistics should be seen as the start of the development of environment statistics in the countries participating in the RETA. As the collection of environment statistics becomes part of regular data collecting activities, both the statisticians and environment experts will need to address all the outstanding methodological issues. When that happens, a contribution will have been made to the development of environment statistics in the countries along with other statistics.