Research and Management Techniques for the Conservation of Sea Turtles
K. L. Eckert, K. A. Bjorndal, F. A. Abreu-Grobois, M. Donnelly (Editors)
IUCN/SSC Marine Turtle Specialist Group Publication No. 4, 1999

Databases

Raquel Briseño-Dueñas and F. Alberto Abreu-Grobois
Banco de Información sobre Tortugas Marinas (BITMAR), Estación Mazatlán, Instituto de Ciencias del Mar y Limnología UNAM, Apdo. Postal 811, Mazatlán, Sinaloa México 82000; Tel: +52 (69) 852848; Fax: +52 (69) 82613; email: raquel@servidor.unam.mx, abreu@ola.icmyl.unam.mx

Because of sea turtle life cycle characteristics, trends in the population dynamics of sea turtles are analyzed from essential data accumulated over many years, typically for more than a decade and, ideally, should be based on information for all life stages. Even so, most databases are comprised of data gathered from nesting females; that is, by beach monitoring programs. These data, when properly collated, can still provide a meaningful and representative evaluation of population dynamics. The scenario becomes progressively more complicated, however, when individuals comprising a specific management unit nest at multiple beaches, necessitating monitoring at more than one site. To obtain an accurate representation of population dynamics in this case, data need to be shared and integrated, rapidly and efficiently, among localities and sometimes across national borders. Raw data by themselves have little value. Only through proper collecting, organizing, processing and presentation do they become meaningful. Properly managed and structured databases enhance the efficiency of information archival and transfer.

This chapter describes a model database system for GIS-compatible information. The model is applicable for the management of long-term data from a single sea turtle project and, if adopted by a multi-rookery program with information exchange between data-gathering programs and management agents, the model can bolster integrative collaboration within or among nations. In a multi-program application, a high degree of standardization in methodology, terminology, etc. is required to permit the exchange and comparison of data across space and time. The design presented here borrows extensively from a system employed for Indo-Pacific information management (Limpus, 1995), as well as from a national sea turtle database in México (Briseño-Dueñas and Abreu-Grobois, 1994).

In general, managing sea turtle information through a database should enable: (1) updated and sufficient information for the purpose of monitoring the conservation status of individual management units (breeding stocks) and assessing conservation and management programs; (2) long-term storage and retrieval of data related to sea turtle biology; (3) rapid transfer and exchange of standardized information among research and/or monitoring programs; and (4) the accumulation of long time-series of population parameters useful in robust analyses of population dynamics. There are two organizational levels to be considered: the organization of participants (especially if the database will manage data from more than one project) and the organization of the data, or database structure.

Organizing a Database

Organizing Participants

For regional or multinational applications, participants will be drawn from a variety of organizations and jurisdictions. Such heterogeneity in the membership will undoubtedly require delicate prior agreements on the role each party will have in the management of the database. Agreement must be reached on the rights and obligations of each participant, the custodianship of the database, and proper standards for the complete information cycle. Most likely there will be government agencies, universities, NGOs, and (sometimes) private enterprises involved in data-gathering activities. Despite issues of general interest *(e.g.,* recovery of sea turtle populations, maintenance of quality in the data), short-term commitments and requirements may vary widely among program partners. Management authorities, for example, will need timely periodical reports to assess the success of their programs; research scientists will require high standards in data quality and assurances of intellectual property protection. Most everyone will want to safeguard the proper acknowledgment of data authors.

Issues to Consider

Individual and collective legal requirements must be reconciled, particularly when the project includes a complex array of participants. In such a case, establishing a coordinating group (sometimes referred to as a "hub") is useful and should be comprised of representatives entrusted with decisions related to the management process. The hub might also handle inquiries, archive key documents and publications for distribution, and maintain backup copies of selected data sets for increased security. Further issues to consider include: (1) rights of access to the data, both by users participating in the database development project as well as by outsiders; (2) protection of intellectual rights and proper acknowledgment; (3) appropriate uses of the data; (4) custodianship; and (5) validation standards for data. Control over data access in modern computer software is relatively easy to achieve. In multi-user contexts, an array of user names and passwords can be designed to give access solely to accredited parties. If desired, each user can be controlled by differing degrees of "privileges" in data access. For example, access restrictions (designated as full, limited, or none) can be implemented, depending on the user, to whole data sets or even to individual records.

Intellectual Property Rights

Perhaps the single most delicate point in a database involving many organizations, particularly one in which academic institutions are involved, is the need to ensure adequate regard for intellectual property rights. Clear guidelines on appropriate use are essential, particularly when data gathered by one party are required by another. In the case of data sets released for general consultation by accredited users, it may be enough to guarantee that the source is acknowledged whenever the data are used. But in cases where participants provide data that are considered sensitive, safeguards are prudent; in particular, due regard must be afforded to the privileges of academic researchers to publish original findings. One solution is to specify access constraints over a specified time period and dictated by the data provider for specific records or data sets. The constraints (data present but not available for consultation) could be sustained over a length of time *(e.g.,* 1-2 yr.) considered by the parties reasonable.

Custodianship

Implementing custodianship alleviates several potential problems, and helps to ensure stability and quality in the database. In the absence of designated custodianship, important management tasks might be duplicated, neglected or omitted. Custodianship entails a strong commitment to guarantee various aspects, including (from WCMC, 1996): (1) advising users on potential database use(s), including permitted and forbidden usage (uncertainties or ambiguities could be pointed out on specific data sets); (2) ensuring that publication, information products and all outputs derived from the database acknowledge data sources and protect intellectual property; (3) coordinating the coding of parameters applied to major referential variables *(e.g.,* project sites, organizations, personnel, geographic reference grid); (4) ensuring that the database is up-to-date, adequately documented, and maintained in such a way as to be accessible to users; (5) undertaking periodic updating, safe backups and adequate virus protection; and (6) proposing occasional changes to structure and content as needs arise. Monitoring of tagged turtles produces information needing special consideration since the data are useful to both the tagging and recapturing parties. The custodian's responsibilities should include the facilitation of this access, again with respect to intellectual property rights.

Designation of custodianship may not be easy. Legal considerations may dictate that the custodianship be awarded to a government agency, yet greater continuity and quality can sometimes be achieved if the task is given to a reputable academic organization. In general, custodianship should be conferred on the organization most familiar with the history, management characteristics, and potential uses of the database. Broad consensus may be necessary when several groups contest the claim. Principally, the custodian group must be technically capable, inspire confidence in the users, and have an acceptable long-term stability. Organizations with stable financial resources and a prior track record in the field are good prospects as long as they are also considered to be impartial (lacking conflict of interest).

Organizing Data

The organization or structure of the database to be adopted should both contain enough fields to hold information on key parameters of sea turtle population dynamics and be distributed in such a way as to avoid redundancy of space usage, while facilitating searches (or queries) and data retrieval. Further, two levels of sea turtle monitoring capability must be recognized and provided for in a database of broad applicability. Programs with sufficient expertise and resources generate high-resolution data, where information on individual turtles is available through nesting or capture/tagging monitoring. Programs characterized by more limited resources oftentimes do not monitor individual turtles; nevertheless, basic data are collected on key parameters from global, whole-nesting beach surveys. The majority of sea turtle monitoring programs fall somewhere between these two types as it is often impossible in practice to locate, tag and monitor even the majority of turtles nesting on a single beach, particularly where the population is abundant.

Validation

Validating information from field sheets is essential to database quality. Unfortunately, in many database applications this aspect is weakly observed. For efficient data analysis, formats for each data attribute and input must be adhered to strictly and with total consistency. Use of numerical coding for repetitive data helps to avoid typographic errors. Validation should be done at the point of origin, before transferring to a global database. Although some programs provide automatic checking for the simplest mistakes *(e.g.*, verifying that all required data fields are filled, cross-checking locations given against a master catalog, ensuring that repetitive and constant data are in corresponding fields), only direct revision on hard copies of entered data should be considered reliable for the detection of error(s). Temporary files can be implemented to hold data that only after full validation get transferred into master files.

Compatibility, Software and Hardware

Above all, compatibility in the process of data exchange needs to be ensured so all parties have access to the information. While standardized methods for data gathering are basic, compatibility in software and hardware is also important. Even though modern software has facilitated converting data from one format to another, compatibility is far from perfect between products of different manufacturers. Choice of software is thus critical and, particularly for multi-user applications, selecting a single product is highly recommended, particularly when further analysis of the data with sister-applications will be sought.

There are ample choices in database management software. Consideration should be given to: (1) volume capacity (is it capable of managing the volume of data and number of users?); (2) expansion (can it cope with future increases in user volume? does it contain easy to learn facilities for developing application; *e.g.*, entry screens, querying tools, reporting facilities?); (3) is the product likely to continue being supported and enhanced?; (4) does the program have sufficient experience for the development of applications and maintenance of the software? Most often, a choice of software will hinge between selecting a popular package affording ease of use, a quick learning curve, and sufficient power for modest single project applications (the simplest representatives cost

Table 1. Structure for NESTING BEACHES CATALOG (considered essential)

elements of table	fields	notes
	— each record to contain descriptive data for a single nesting site —	
· ***date***	- data entry date	
· ***descriptors of site***	- ***beach code*** - beach name, abbreviation - other local names - total size - actual length protected	preferably defined by national or international coding system
· ***location***	- latitude, longitude - country, state, municipality - reference landmark (natural or town)	
· ***source of data***	- name of person, affiliation, ***personal code, institutional code***	coding system could **link** to national/international international database of personnel and institutions
· ***others***	- other significant parameters	ecosystems, land use/tenure, management authority, significant developments, disturbance factors
· ***notes***	- complementary observations	

Table 2. Structure for ANNUAL BEACH SURVEY (considered essential)

elements of table	fields	notes
	— each record to contain data from a single nesting site for a given year —	
· ***date***	- date of data entry	
· ***descriptor of locality***	- ***beach code***	**links** with NESTING BEACHES CATALOG
· ***year/season***	- ***nesting season***	may prefer to use format such as 95/95 or 95/96 to allow for nestings spanning more than one calendar year
· ***locality/season***	- ***beach-season code*** (concatenated)	**links** with ANNUAL NESTINGS BY SPECIES TABLE
· ***survey coverage***	- extent of beach protected during given season	permits estimation of variations in survey coverage
· ***responsibility***	- management authority or organization (govt, international instit. university, etc.)	
· ***source of data***	- name of person, affiliation, ***personal code, institutional code, literature citation code***	coding system could **link** to national/international database of personnel and institutions or bibliography
· ***notes***	- complementary observations	

Table 3. Structure for NESTING Table (high-resolution)

elements of table	fields	notes
	— each record to contain observations from a single nest —	
· ***date***	- date nest laid	
· ***event registration number***	- ***registration number***	id for observation, corresponds to reg. number on field data sheets; use incremental numbers, starting every year
· ***event code***	- composite ***event code*** (made from concatenating site+season+registration number codes)	unique record id **links** with CAPTURE TABLE if nester seen; facilitates ordering and flagging of records by site of origin
· ***clutch data***	- nest number - total number of eggs laid - complete or partial clutch - which clutch of the season - number of eggs incubated - yolkless eggs - multiyolked eggs - quantification of other descriptors of embryo (e.g., partial development)	
· ***egg data***	- egg diameter, weight	best if summarized as mean, std. dev., range and sample size; can use individual measurements but will need separate tables
· ***hatching data***	- number of eggs hatched - number of dead hatchlings - number of deformed hatchlings - number of hatchlings released - number of females	can derive estimates of hatching success from these these data sex ratio can be estimated
· ***nest data***	- depth to top egg, to bottom - nest location in beach	
· ***fate of clutch***	- relocation code	final incubation site (e.g. *in situ*, beach hatchery, incubation house)
· ***applicable restrictions***	- apply restrictions yes/no - specifications of restrictions - length of time data to remain restricted	
· ***data source***	- name of person, affiliation responsible for the data ***personal code, institutional code, literature citation code***	coding system could **link** to national/international database of personnel and institutions, or bibliography if data are obtained from publications
· ***notes***	- complementary observations	

US$ 100-1,500; *e.g.*, Access, Paradox, dBase) and a sophisticated system with specialized database engines designed for efficient simultaneous multi-user, multi-platform and rapid remote-access by 50 users or more. At the latter end, databases can contain more than 10^6 records and Unix-based platforms are recommended (the most expensive software can cost >US$ 10,000; *e.g.*, SQL Server, Oracle).

The selection of a computer type *(e.g.*, PC-IBM, Macintosh, UNIX-based) has almost become a moot question, as manufacturers steer towards greater integrating capacities, particularly between and within networks interconnected locally or remotely. When a large database project is envisaged, only specialized servers should be considered. Connectivity to enable rapid exchange of data between individuals and organizations is now possible and economical with current electronic communication among local area networks (LAN) or between remote stations employing the Internet (either e-mail for batch queries, or interactive consultations through a Wide World Web interface). All modern software have capability for remote data searching via any of these. In all cases, expert advice is recommended particularly since integration of software and the development of queries and further analysis of data will require some degree of programming.

Documenting the Database

Often existing information is underutilized, largely because its location, content and applications are unknown. To avoid this, databases should include adequate documentation, providing descriptions of the structure, name, format and fields (data dictionaries) together with information about location and any policies regarding data access. As a whole, these metadata ("data on data") can clarify to users and outsiders the content, functions, and management of a database. Collaborative work and further consultation for management practices are enhanced; sharing, linking and improving existing databases are also facilitated.

Table 4. Structure for ANNUAL NESTINGS BY SPECIES Table (considered essential)

elements of table	fields	notes
—— *each record to contain data for a single species nesting at site for a given year* ——		
· ***date***	- date of data entry	
· ***locality/season***	- ***beach-season code*** (concatenated)	**links** with ANNUAL BEACH SURVEY TABLE
· ***species***	- species, ***species code*** - management unit nesting here, ***MU code***	species and management units may be given numerical codes to facilitate queries; **links** with species and MU tables (optional)
· ***species season census data***	- time span of survey, dates - estimated time span of nesting activity, dates	permits estimation of variations in survey coverage from season to season
	- total females counted	stipulate if actual or estimated
	- total dead turtles found	specify raw or processed
	- total nest count	specify raw or processed
	- count of nests destroyed	specify raw or processed
	- count of eggs protected	poached nests, nests destroyed by natural causes
	- estimate of eggs lost	
	- count of hatchlings released	
	- estimation of total females, nests, eggs	
	- methodology code	extrapolation to full extension in cases when only partial surveys are possible; specify which
· ***annual rookery size***	- estimated size (females, nests) - methodology of derivation	value should include conf. limits statistical basis
· ***conservation significance***	- significance coding by species	rating for each species: e.g., rookery size relative to global size of management unit
· ***source of data***	- name of person, affiliation, ***personal code, institutional code, literature citation code***	coding system could **link** to national/international database of personnel and institutions or bibliography
· ***notes***	- complementary observations	

Database Application

The database model presented here employs a relational design, with basic entities (tables) holding data in a set of rows (the records) and columns (the fields). As the name indicates, tables are "related" or "linked" operationally through common fields. For simplicity, field formats (name, type, size) are not specified and descriptions of content appear in a generic manner that will require final assignment by someone experienced with database construction (assignments should suit individual needs). Copies of a documented working database adopting the same basic design can be obtained from the authors by request to exemplify a working application.

The parameters considered essential for sea turtle conservation goals are distributed among separate

Table 5. Structure for CAPTURE Table (high-resolution)

elements of table	fields	notes
	—— each record to contain observations on a single turtle ——	
· ***primary tag number***	- ***number of primary tag***	**links** with TAGS CATALOG; if this is a recaptured turtle, primary tag number needs to be confirmed from tag catalog; if secondary tag seen, entry should be replaced with primary tag number
· ***tag***	- presence/absence of tag - status (tag applied for first time, recapture, old tag replaced) - position of tag - presence/absence of tag scar(s)	tag status should be coded to enable flagging
· ***date***	- date of observation	
· ***event registration number***	- ***registration number***	id for observation, from field data sheet. If turtle laid nest, will be the same number as in NESTING TABLE (**link**)
· ***descriptors of turtle***	- ***composite event code*** (concatenation of site code+year+registration number) - maturity - sex - species	same as in NESTING TABLE (**link**) if capture is in nesting beach, otherwise parallel non-nesting capture site coding to be used
· ***turtle measurements***	- carapace length (curved or straight) - carapace width (curved or straight) - turtle weight - tail lengths (from carapace, plastron, vent) - head lengths (length, width)	type of measurement needs to be standardized (can use more than one but in separate fields)
· ***primary activity of turtle***	- activity	in the context of the turtle's life cycle
· ***capture method***	- capture method	
· ***health and condition of turtle***	- health, condition, stranded	several fields may be necessary and could include information on stranded turtles
· ***additional experimental actions on turtle***	- experiments	optional, related to specific projects
· ***location***	- latitude, longitude if turtle caught in open sea, or, sector in a national/regional grid system - ***beach code*** if found on nesting site	**link** to NESTING BEACHES CATALOG
· ***reproductive history of turtle***	- remigration interval - total number of clutches for the breeding season	
· ***applicable restrictions***	- apply restrictions yes/no - specifications of restrictions - length of time data to remain restricted	
· ***data source***	- person, affiliation responsible for data, ***personal code, institutional code, literature citation code***	coding system could **link** to national/international database of personnel and institutions, or bibliography if data are obtained from publications
· ***notes***	- complementary observations	

Table 6. Structure for TAGS CATALOG (high-resolution)

elements of table	fields	notes
	—— *each record to contain data for single tag* ——-	
· ***primary tag***	- ***number of primary tag***	this is the first tag number applied and should be used to identify turtle throughout its life; when additional tags are applied, the number is repeated for cross referencing; may be convenient to separate alpha-numeric prefix and numeric portions of number into separate fields
	- "return to" data (institution/address to be contacted) - tag type and/or material - tag position on body of turtle	
· ***secondary tag***	- number of applied tag - "return to" data (institution, address to be contacted) - tag type and/or material - tag position on body of turtle	replacement or multiple additional tags
· ***date***	- date of data entry	
· ***special tags***	- tag number, position	description of type
· ***source of data***	- name of person, affiliation responsible for the data ***personal code, institutional code***	coding system could link to national/international database of personnel and institutions
· ***observation code***	- ***event code***	**links** with CAPTURE TABLE
· ***notes***	- complementary observations	

tables following a logical, thematic organization. This modular design ultimately avoids redundancy while facilitating use of some of the tables depending on the immediate needs (or limitations) of the project and extending into other parameters as capacities increase. Parameters considered essential are contained in fields in the NESTING BEACHES CATALOG, ANNUAL BEACH SURVEY, and ANNUAL NESTINGS BY SPECIES tables (Tables 1, 2, and 4, respectively). Minimally, a monitoring program should generate data for these, which can be complemented by data found in historical or current publications. As capabilities are extended, rigorously collected high resolution data *(e.g.,* based on monitoring of individual, tagged turtles) may be added to the database by implementing the remaining NESTING, CAPTURE, and TAGS CATALOG tables (Tables 3, 5, and 6, respectively).

The annual registers should be viewed as important means with which to summarize the major parameters that facilitate an evaluation of conservation results on a beach-by-beach basis, incorporating important pieces of information on factors such as beach survey coverage and mortality (poaching, strandings, natural disasters, etc.) that are normally assessed on an overall basis. For example, the compilation of data through use of the ANNUAL tables can be used to assess conservation status in connection with recovery goals or benchmarks *(e.g.,* target number of nests per season).

Literature Cited

Briseño-Dueñas, R. and F. A. Abreu-Grobois. 1994. Las Tortugas Marinas y sus Playas de Anidación. Informe Final del Proyecto UNAM-CONABIO PO66. Octubre de 1994. 57 pp.

Limpus, C. 1995. Conservation of Marine Turtles in the Indo-Pacific Region. 6. Indo-Pacific Marine Turtle Database. Final Report. Conservation Strategy Branch, Queensland Department of Environment and Heritage. August 1995. 26 pp.

World Conservation Monitoring Centre. 1996. Guide to Information Management in the Context of the Convention on Biological Diversity. U.N. Environment Programme, Nairobi, Kenya. (available from http://www.wcmc.org.uk/capacity_building/docs.html)