ADDRESSING UNCERTAINTY THROUGH DATA QUALITY ISSUES IN ENVIRONMENTAL ASSESSMENTS TO SUPPORT IMPROVED DECISION MAKING

Maria Pantazi[1] ,Vassiliki Vassilopoulou[1], Yiannis Issaris[1], Athina Kokkali[1], Maria Salomidi[1], Alexandros Frantzis[2], Jacob Fric[3], Aliki Panou[4]

1 Hellenic Centre for Marine Research, Ag. Kosmas, Greece.
2 Pelagos Cetacean Research Institute, Vouliagmeni, Greece
3 Hellenic Ornithological Society, Athens, Greece
4 Archipelagos – environment and development, Kifissia, Greece,

Introduction

Acknowledgement and management of data quality assessment in environmental science for policy has become a key factor during the last decades. In the framework of the MESMA (Monitoring and Evaluation of Spatially Managed Areas) FP7 project, an evaluative account of different qualitative aspects of the data (Funtowicz & Ravetz, 1990) describing the natural and socioeconomic components that occur in the region of the Greek case study (Inner Ionian Archipelago and the adjacent gulfs) has been attempted. This process can shed light on the uncertainty inherent in the data and should be considered particularly in cases that will be used in decision-making. Uncertainty assessment has been implemented to a broad range of research fields such as maritime management, environmental and biological modelling, climate change modelling and water management modelling, adding further credibility to the studies' results (Maxim & van der Sluijs, 2011).

Data Quality Assessment

A qualitative assessment of semi-quantitative criteria based on a pedigree matrix that describes those aspects of data quality influencing the reliability of the overall result, has been developed as suggested by Pedersen Weidema & Wesnaes (1996). In our case, the specific pedigree matrix has been modified in order to help specify the reliability, completeness, temporal correlation, geographical correlation and data collection process quality of the ecosystem components under study (Figure 1). Each characteristic is divided into five quality levels with a score between 1 (high quality) and 5 (low quality). The "reliability" indicator relates to the sources, the acquisition methods and verification procedures used to obtain the data. The "completeness" indicator relates to the statistical properties of the data. The "temporal correlation" indicator represents the time correlation between the year of study and the year the data were obtained. The "geographical correlation" indicator illustrates the geographical correlation between the defined area and the location of origin of the data used. Finally, the "data collection process quality" indicator refers to the collection process of the data.

Indicator Score	1	2	3	4	5
Reliability	Measured Data	Verified data partly based on assumptions	Non-verified data partly based on qualified estimates	Qualified estimate (e.g. by scientific expert)	Non-qualified estimate
Completeness	Representative data from all sites relevant for the study area considered over an adequate period to even out normal fluctuations	Representative data from >50% of the sites relevant for the study area considered over an adequate period to even out normal fluctuations	Representative data from only some sites (<50%) relevant for the study area considered OR >50% of sites but from shorter periods	Representative data from only one site relevant for the study area considered OR some sites but from shorter periods	Representa-tiveness unknown or data from a small number of sites AND/ OR from shorter periods
Temporal correlation	Less than 3 years of difference to year of study	Less than 6 years difference	Less than 10 years difference	Less than 15 years difference	Age of data unknown or more than 15 years of difference
Geographical correlation	Data from area under study	Average data from larger area in which the area under study is included	Data from area with similar environmental conditions	Data from area with slightly similar environmental conditions	Data from unknown area or area with very different environmental conditions
Data collection process quality	Data from targeted research conducted by the team involved in the case study	Data from targeted research conducted by other teams not involved in the case study	Data from targeted research conducted with different methodologies	Data from common research conducted with a standard methodology	Data from common research conducted with different methodologies

Figure 1: Modified pedigree matrix used for the assessment of the semi-quantitative indicators in our case study

Results

In an attempt to visualize the qualitative information gained from the implementation of the pedigree matrix, radar diagrams were produced (Figure 2), presenting each indicator's score for every ecosystem component. Hence, certain ecosystem and socioeconomic components (coralligenous and deep corals, fan mussel, gold coral and small-scale fisheries) exhibited low quality assurance (high uncertainty) mainly due to the lack of completeness or the fragmented data collection process. On the other hand, other ecosystem and socioeconomic components (loggerhead turtle nesting beaches, trawlers, purse seiners, industry, and tourism (blue flags and diving centres)) having higher data quality exhibited high quality assurance and therefore low uncertainty. The above should be also taken into account while designing appropriate monitoring programmes aiming to fill the identified gaps.

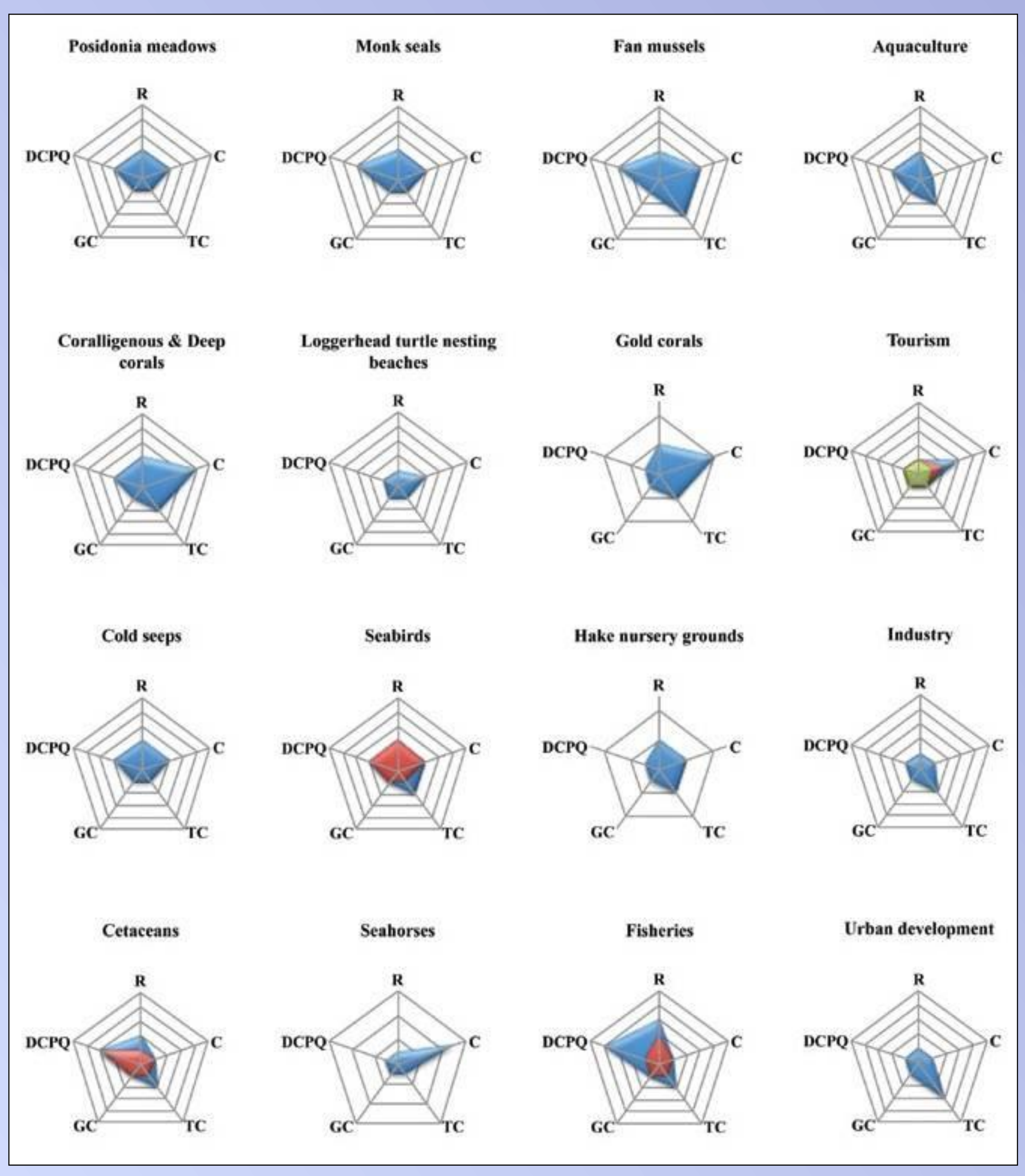

Figure 2: Radar diagrams for the ecosystem components, showing the scores for each indicator (R: Reliability, C: Completeness, TC: Temporal correlation, GC: Geographical correlation and DCPQ: Data collection process quality). For the ecosystem component: Tourism, the blue pentagon corresponds to marinas, the red pentagon to diving centres and the green to blue flags. For the ecosystem component: Seabirds, the red pentagon corresponds to the species C. diomedea and the blue to the species P. aristotelis desmarestii. For the ecosystem component: Cetaceans, the red pentagon corresponds to the species P. catodon, Z. cavirostris, T. truncatus, & S. coeruleoalba and the blue to the species D. delphis. For the ecosystem component: Fisheries, the blue pentagon corresponds to the coastal fleet and the red corresponds to the purse seiners & trawlers fleet.

Conclusions

The results of the data quality assessment indicated that the sampling effort was not evenly distributed throughout the study area; thus for many ecosystem components the accuracy of assessments could be hindered by lack of completeness in the data collection process. In terms of data reliability, specifically for small-scale fisheries, since, due to the low data availability, expert judgment was applied in the analysis process, further sources of bias were introduced. However, it should be noted that the geographical and temporal correlation of the overall ecosystem components exhibited high quality assurance (Issaris *et al*., 2012).

References

1. Maxim, L. & Van der Sluijs, J.P., 2011. *Quality in environmental science for policy : assessing uncertainty as component of policy analysis.* Environmental Science & Policy, 14 (4): 482-492.

2. Funtowicz, S.O. & Ravetz, J.R. (1990) *Uncertainty and quality in science for policy.* Dordrecht, Kluwer Academic Publishers, 244 pp.

3. Weidema, B.P. & Wesnæs, M.S. (1996) *Data quality management for life cycle inventories - an example of using data quality indicators*. Journal of Cleaner Production, 4 (3-4): 167-174.

4. Issaris Y, Katsanevakis S, Pantazi M, Vassilopoulou V, Panayotidis P, Kavadas S, Kokkali A, Salomidi M, Frantzis A, Panou A, Damalas D, Klaoudatos D, Sakellariou D, Drakopoulou V, Kyriakidou C, Maina I, Fric J, Smith C, Giakoumi S, Karris G (2012) *Ecological mapping and data quality assessment for the needs of ecosystem-based marine spatial management: Case study Greek Ionian Sea and the adjacent gulfs.* Mediterranean Marine Science 13:297-311