



Estimating Impacts of Environmental Interventions in Monitoring Programs Requires Conceptual Data Models and Robust Statistical Processing

Ladislav Dušek, Jana Klánová, Jiří Jarkovský, Jakub Gregor, Richard Hůlek,
Ivan Holoubek, Jiří Hřebíček

► To cite this version:

Ladislav Dušek, Jana Klánová, Jiří Jarkovský, Jakub Gregor, Richard Hůlek, et al.. Estimating Impacts of Environmental Interventions in Monitoring Programs Requires Conceptual Data Models and Robust Statistical Processing. 10th International Symposium on Environmental Software Systems (ISESS), Oct 2013, Neusiedl am See, Austria. pp.204-221, 10.1007/978-3-642-41151-9_20. hal-01457450

HAL Id: hal-01457450

<https://inria.hal.science/hal-01457450>

Submitted on 6 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Position paper

Estimating impacts of environmental interventions in monitoring programs requires conceptual data models and robust statistical processing

Ladislav Dušek^{1,2}, Jana Klánová², Jiří Jarkovský¹, Jakub Gregor^{1,2},
Richard Hůlek², Ivan Holoubek², and Jiří Hřebíček¹

¹Institute of Biostatistics and Analyses, Masaryk University,
Kamenice 126/3, 625 00 Brno, Czech Republic

{dusek, jarkovsky, gregor, hulek, hrebicek}@iba.muni.cz

²Research Centre for Toxic Compounds in the Environment, Masaryk University,
Kamenice 753/5 (A29), 625 00 Brno, Czech Republic
{klanova, holoubek}@recetox.muni.cz

Abstract. This paper discusses main problems associated with evaluation of performance and impact of long-term environmental programs. Lack of data standards, incompleteness of archived datasets and insufficient statistical power were identified as main limits in functionality of monitoring networks. To avoid these failures, environmental programs should be designed with inception to incorporate data management as their integral part. Especially in global programs, local and regional data managers should invest significant proportion of their effort to handle documentation in terms of standardized coding, data formats, metadata coding and consistency of records over time. Up-to-date trends in building knowledge-based infrastructures are illustrated using example of monitoring of atmospheric pollution by persistent organic pollutants (POPs). Conceptual model usable to facilitate the integration and analysis of data on POPs concentrations is introduced with its multilayer hierarchy of entities (POPs as nomenclature classes, couples “observation – measurement” as content classes). Robust set of statistical methods for processing of time series of concentration data is discussed from the viewpoint of practical implementation within running monitoring programs. It consists of the following components: baseline pollution estimates, uncertainty analyses, spatial extrapolations, effect size estimates, time trend identification and quantification. Development of tools supporting standardized environmental data management is rapidly expanding field of science which results in the following challenges for applied informatics and statistics: long-term sustainability of information systems, data-related metadata coding and archiving, tools for automated integration and reporting of data.

Keywords: environmental monitoring, persistent organic pollutants, data model, data standards, statistics.

1 Introduction – assessing impact of environmental programs

For the past several decades, there has been considerable effort to establish representative environmental monitoring networks as a response to increasing pollution by hazardous compounds released from various anthropogenic sources. The effort has resulted in more or less specialized monitoring programs, designed to cover various environmental matrices. Recently adopted international environmental policies make it imperative for most of the developed countries to upgrade their national monitoring programs, to keep them sustainable and to share exported data in continental and global level. However, the evaluation of various policies launching environmental interventions still remains a rather tentative domain, which is often reasoned as a consequence of the complexity of the problem, multiplicity of stakeholders to be addressed or by prevalence of multilevel uncertainties [1-2].

Nevertheless, increasing spatial and temporal scales of environmental programs strongly stimulate needs to compare their performance and to evaluate their reachable cost/benefit ratio. Special emphasis is placed on standards related to the quality of source data which form the real base of any functional monitoring [3-4]. Due to recently sharply growing availability of data from networks covering the whole continents or entire globe, the task of analyses performed at larger scales has become feasible, among others due to affordable computational costs and capacity of information systems [5-6]. Many complex methods, models and scenarios have been employed in the assessment of outcomes of environmental programs [7-11]; however, the process of evaluation still basically depends on applied stochastic procedures. From this viewpoint, we can distinguish two principal approaches in the evaluation of environmental programs, regardless of their content and focus, which corresponds to the common classification of statistical techniques:

1. Descriptive evaluation, typically aimed to extract unbiased estimates of proper summary statistics, mostly generated from varying or incomplete database records.
2. Inferential evaluation which is concerned with testing hypotheses and with making predictions or other inferences about examined environment. In this case, the main topic is generalization of outcomes from monitoring programs to a larger scale.

Although both approaches play important role, only inferential methodology can contribute to the impact assessment of environmental interventions:

1. Spatial comparison and/or comparability of different monitoring programs or data sources. Information on inter-regional or inter-continental differences in pollution is helpful in identification of contamination sources.
2. Assessment of time-related trends in pollution which can be attributed to a specific source (in case of increasing trend) or to a particular effective prevention or intervention (in case of decreasing trend). Statistically significant decreasing trend in pollution can be accepted as proof of impact of the examined program.

Impact assessment is becoming obligatory component of current environmental programs, which, however, must be intrinsically incorporated in the study plan from

their very beginning. Otherwise we must solve many problems like selection bias, insufficient comparability of sites or impact of historical events which coincided with the evaluated program and thereby provided an alternative explanation for the observed trends. To avoid these problems we need properly designed monitoring plans which can transform the collation of data into robust information systems enabling standardized comparative analyses of sites and regions. The main goals of this paper are to outline statistical methods useful in such analyses, to discuss robust approaches which can overcome weaknesses associated with historical environmental data and to document rules to be adopted in prospectively planned monitoring studies. Main recommendations will be illustrated using example of monitoring of air pollution by persistent organic pollutants (POPs).

2 Current limits in the evaluation of environmental programs

One of the main causes of difficulties in the evaluation of environmental programs is inadequately designed databases which then come into a conflict with desired end-points. These contradictions form an essential basis of many problems that cannot be solved retrospectively once the data are obtained in inadequate form. To avoid it, proper standards including statistical power planning should be implemented in obligatory guidelines of any “evidence-based” monitoring program. This results in challenges for current development in this field.

Most problems limiting information value of the data arise from the heterogeneity of input data ranging from laboratory biotests to multilevel epidemiologic observations. There is no “universal” truth in the outcomes of various environmental programs. Progress increasingly requires standardized access to multi-disciplinary information resources, including chemical, geological, meteorological, epidemiologic and demographic data. Each broadly ranged ecological study must adopt both following scenarios [12-13]:

1. retrospective exploitation of data sources and their description in discovery process;
2. prospective arrangement enabling effective electronic data capture in future.

Another source of problem is proper selection of territories where the monitoring is conducted. Especially for long-term programs it might be difficult to keep the network sustainable in sites with stable environmental conditions. In this sense, the information sustainability of the programs means that they are able to gather data in continuously comparable sources (sites), matched in a paired (repeated measures) design. The network cannot randomly assign data obtained from individual sites in past to recent outcomes from different sites or to some randomly selected control (background) sites. Growing number of studies is not accompanied by adequate progress in information background and prospectively optimized plans [14-15]. It inevitably results in publishing of non-consistent outcomes with ad hoc data management support. To discover such broadly heterogeneous data we need consensus on data and metadata standardization, but this alone is not enough. We need sufficiently compre-

hensive conceptual models, advancing development of formal ontologies over environmental data capture systems.

Although there are some usable standardizing concepts already published (Ecological Metadata Language) [16-17], they are not extensively used in practice. Environmental data collection is still subject to research in the informatics field. Especially in global monitoring networks, we need standardized documentation of data in the concept of metadata. Metadata may contain handling instructions for data, such as details of its format, information about its quality, and information required by the user to determine its accessibility for different modes of use.

To conclude, widespread support for ontology-based approaches is not implemented in the field of ecological risk assessment [18]. However, interest in developing ontologies has been growing, because new synthetic environmental analyses increasingly rely on access to a broad range of cross-disciplinary data sources and monitoring studies. The effective system should encompass not only structure and content of such data repositories, but also hierarchical architecture and mutual relationships among components [19–21].

3 Persistent organic pollutants (POPs) as model

To illustrate the role of standardized data management in environmental studies, we took persistent organic pollutants (POPs) as a proper model namely due to the following reasons:

- POPs have become recently intensively studied due to their properties which represent remarkable risk for ecosystems and human population (persistence, bioaccumulation, carcinogenicity, genotoxicity).
- Clear nomenclature of POPs serves as suitable classification system which facilitates development of formal concepts based on well defined entities.
- POPs serve as a proper model for global monitoring studies because these pollutants are in focused attention of several widespread monitoring programs which mostly deal with air sampling as driven by the Global Monitoring Plan (GMP) of the Stockholm Convention (SC; negotiated under the auspices of the United Nations Environmental Program, UNEP) [22]. Any contribution to the standardization of these monitoring activities can be extensively utilized by the running active programs which have been recently overviewed by Klanova and Harner [23]. Moreover, more than twelve years after the adoption of the Stockholm Convention, the questions on its effectiveness prioritize development of functional IT infrastructures with capability to address both recent and historical time trends in POPs fate in the environment.
- The global framework established by the Stockholm Convention is vital and targets specific groups of POPs [24]. The decisions by the Conferences of the Parties to the Basel, Rotterdam and Stockholm conventions on enhancing cooperation and coordination among the three conventions invite Parties to coordinate their efforts when implementing the three conventions to ensure close cooperation among relevant ministries and programs at the national levels. The Conference of the Parties

initiated „the establishment of arrangements to provide itself with comparable monitoring data on the presence of the [POP] chemicals as well as their regional and global environmental transport” [25].

4 Conceptual data model for environmental monitoring of POPs

Here we proposed simplified conceptual model that should broaden our capability for understanding the validity, content and relevance of the data coming from environmental monitoring (Table 1). The model is based on hierarchically layered architecture providing different levels of classifiers or properties of homologous entities as well as scoring of data origin. The model works with three principle layers:

1. **Entities** defined on the basis of internationally standardized nomenclature systems. The level is linked with classifiers, i.e. given properties extending the nomenclature and filtering homologous groups of entities (in this example: POPs).
2. **Observation – measurement** level and its descriptors, focused namely on time & space coordinates, methodical attributes, measured endpoints, reference benchmarking of their value and validity scoring.
3. **Content** identification describing employed measures, units and precision estimates.

Important principles applied in the model construction are further summarized as follows:

1. Reduction of the number of object properties. When necessary, the set of attributes can be expanded before specifically designed data discovery. Too many object properties cannot be utilized efficiently in retrospective exploitation of the resources.
2. Any relevant data discovery must reflect heterogeneity of experimental and methodical approaches. That is why type of the study or data resource is obligatory attribute among the observation descriptors.
3. A measured value cannot be interpreted without reference to a defined/known measurement standards or reference benchmarks. Both internal reference norms (e.g. self-benchmarking of time series data) and external benchmarks (e.g. background concentration levels or limits, hygienic norms, detection limit of applied method) can be applied.
4. Descriptors of measures must fulfill obligatory measurement standard, i.e., the units, scales and lists of attributes defining origin of the measures (e.g., examined matrix, sampling methods, etc.)

The model template enables the IT tools to integrate various data resources. The relevance of the integration process relies on the ability to determine if two values (studies) are compatible, not only in time and space coordinates. Description of model levels in Table 1 implies multivariate interactions of classifiers extending the nomen-

clature. The most important added value of the model is the capability to determine whether two data sets can be either fully or partially merged or mutually related (correlated) once they are discovered. During the decision process the system undertakes important steps in each level of its architecture:

1. The system must control relevance and compatibility at the taxonomic level (nomenclature) and in space & time localization of data resources
2. Identified data resources must be assessed whether and at what semantic resolution the data are compatible (level of classifiers and/or extending descriptors like type of the study, etc.)

Finally, the measurement standards for the mutually related environmental endpoints have to be controlled for compatibility (units, scales, reference benchmarking, methodical origin).

5 Projection of conceptual model to data standards

The quality of conceptual model determines its utility for assisting in data discovery and information searching. However, the applicability strongly depends on quality of description and content of processed data sets. That is why we should insist on minimized obligatory database components and their descriptors. Minimized data model as standard can be used both retrospectively (scoring of validity of discovered resources) and prospectively (when designing new data capture system). Proposed conceptual model intrinsically encompasses these obligatory items: institution (origin of data), time & space coordinates, type of resource, examined entities (compounds), measures and methodical descriptors (values and units, matrices, methods).

The system allows for any type of reasonable extension; additional properties may be added on demand. However, minimized data standard ensures accessibility of key information namely in the measurement level of the model; i.e. when and where measurements were recorded, who recorded each measurement, the methodology of measurements, study design and aims. In this way the model can improve data visibility for search engines and enables greater levels of automation of common data transformation and integration.

Proper conceptual model also contributes to widely recommended discovery of data according to the concepts they really represent [26-27]. Ontology should represent the knowledge in a domain of interest, defined via the terminology (concepts, nomenclature) used within the domain and the properties and relationships among domain objects [16-17]. This concept is fully implemented in the model proposed here; the nomenclature baseline is extended by selected descriptors with defined dependencies. It is a formal framework for observational studies where we adopted structured approach recognizing key entities (nomenclature classes) in the 1st level and their characteristics (classifiers). Second level consists of measurements and their characteristics, i.e. validity criteria, origin of data, etc. Third level covers content identification, namely values and units, scales.

6 Impact of conceptual model on standardized data management

Implementation of conceptual data models increases the usage value of monitoring programs and facilitates their sustainability. To keep the information system for long-term monitoring activities sustainable, it is necessary to develop some strategy for handling the incorporation of consecutively gathered data into the archive. That is why the description of data, methods and sites in standardized metadata are vitally important – it is the only way how to track the changes in sites, regional networks, etc. [28]. In any evidence-based environmental monitoring, the data structure must be well defined, but flexible enough to reflect a wide range of possible hypotheses. Of course, such system cannot be constructed retrospectively, on demand of running analyses. Baseline standardization through conceptual data model positively impacts upon analytical procedures, namely in the following four fields:

- **Integration of disparate terminology** which facilitates standardized management of metadata, digital libraries and specific tools working with data.
- **Hierarchical structure advances the data analysis.** The conceptual model intrinsically distinguishes hierarchy of levels and descriptors which facilitates implementation of tools focused on data analysis and knowledge mining. The position of nomenclature entities and measurements can be used to denote a wide range of entity characteristics (nominal or ordinal measures). Using the hierarchy of descriptors we can easily decide whether the data are useful for a particular analysis.
- **Conceptual model supports robust comparison of values.** Regarding data analysis, very important attribute is incorporation of measurement level and its characteristics. Validity criteria reflect some precision measures as well as reference values or protocol standards. A measured value cannot be interpreted and analyzed without reference to a defined measurement standard.
- **Stratified analyses and integration of different data sources.** Hierarchical relationships among nomenclature classes and descriptors also potentiate development of automated SW tools for comparison of values using different strata. For example we can summarize concentrations of some pollutant according to site locations because the sites and their load provide context for observation of contamination. The sites and matrices thus provide a context for measurement of POPs exposure. Different summaries can be then interlinked using various time frames. The concept facilitates data integration, reasoned by compatibility of interlinked values.

7 Environmental programs require robust statistical processing

From the viewpoint of statistics, environmental monitoring can be characterized as processing of heterogeneous data leading to probabilistic estimation of some uncertain (prospective approach) or on the other hand relatively certain (retrospective approach) risk event. Environmental studies are complex and require processing of

highly diverse data. Even if we can recently get standardized, structured data fully accessible for analyses, it is often difficult to select the best approach how to

- extract information from data collected in past, mostly prior to standardization;
- realize later analytic steps which assume some specific input or data aggregation from the preceding measurements;
- merge different data sets collected in different times ;
- mutually relate various measured factors.

To illustrate the role of statistical methods typically employed in analyses of long-term monitoring studies we use already mentioned Global Monitoring Plan (GMP) of the Stockholm Convention (SC) which is focused on specific groups of POPs [22], [24-25]. The GMP gathers relevant global data on POPs concentrations in collection campaigns consecutive in time. The volume and diversity of the data generated by the global campaigns necessitated a centralized approach to quality assurance, data management and robust statistical processing. The principal methodical goals are to:

- summarize the GMP data into a comprehensive report, database and on-line reporting tool,
- propose methodology for the effectiveness evaluation of the existing data to be used as baseline estimates against future concentration data which can be measured,
- recommend a strategy for further environmental monitoring efforts to assess temporal and spatial trends in POPs concentrations.

Following six methodical steps were recommended for the statistical processing of the GMP data with respect to objectives of the GMP [24-25] (Table 2):

A. Exploratory statistics:

1. Data and metadata pre-processing.
2. Summary statistics of the baseline concentrations of POPs.
3. Analysis of uncertainty.

B. Inferential statistics:

4. Power analysis.
5. Stochastic identification of the time trends.
6. Quantification of time trends.

7.1 Data and metadata pre-processing

Correct definition of data is an unavoidable prerequisite of all subsequent statistical analyses in any monitoring. Only reliably reported values can be accepted for any spatial or temporal comparisons. The prior evaluation procedures guarantee comparability of different samples, especially from the viewpoint of site type, matrix, sampling method, time span and sampling frequency. Heterogeneity in these factors might

dramatically increase the uncertainty in the final outcomes. The pre-processing procedures also limit the impact of uncontrolled covariates and thus reduce the risk of a false detection of a trend or the risk of neglecting truly significant changes.

Metadata as “data describing data” represent extremely important standardizing component, mainly in long-term programs where historical datasets have been planned to be merged with the recent ones. Each data set should be supplied with the following minimum metadata profile:

- site identification (coordinates, name, ID, responsible data manager)
- site typology (background vs. impacted, type of anthropogenic influence, etc)
- information value of the program (type: long-term, occasional, ad hoc study, etc.)
- information value of data (primary data vs. aggregated values, type of aggregation)
- sampling strategy and frequency, applied design
- quality control (scoring of usability of records)

7.2 Summary statistics of the baseline concentrations of POPs

Annually aggregated POPs concentrations calculated as arithmetic means of the primary values were recommended for quantitative analyses. The quantitative estimates of concentration values in a specific historical period are called baseline estimates. Variability of baseline values can be quantified on a local, regional or global level, merging appropriate data sets. Pooling of underlying data (both primary and aggregated), however, has to be supported by an uncertainty analysis (see Chapter 7.3). Non-parametric tests like Mann-Kendall U test or Kruskal-Wallis test are recommended for the inter-regional comparison of the POPs concentrations. Parametric tests like ANOVA models or analysis of covariance can be applied only after effective normalizing transformation of the concentration estimates.

7.3 Analysis of uncertainty

The decision makers working with the outcomes of POPs monitoring meet widely varying types of information, each type being associated with its own uncertainties. Local occasional measurements reveal different distributional profiles than long-term, regularly sampled time series. Although the POPs concentrations are mostly collected as a part of some large monitoring network, they must be correctly spatially referenced, since they apply to a specific territory (site). The uncertainty measures therefore mostly reflect spatial and time-related variability [29].

As data reported to the global monitoring networks are typically generated by a variety of programs from various background sites, they have to be inspected for an intra-regional and inter-regional homogeneity in the annually averaged values. Graphically, the regional variability can be reported as their intra-regional 5th-95th percentile range. Sample distribution functions of the regional samples can be then compared and tested by proper robust method (Kolmogorov-Smirnov test, Kruskal-Wallis test). The same applies for geometric means of the averaged concentrations and their 95% confidence intervals.

Another test of relevance is also an examination whether the information outcome has changed if the dataset was improved or degraded. Multimodal distributional patterns, which often indicate the influence of some “effect modifying factor”, can be proved and distinguished from unimodal sample distribution [30]. The homogeneity should be also assessed in the time trend analyses (i.e. presence and the same direction of the trend changes and the annual difference). A year-to-year difference can be compared among the time-series based on the individual sites. Such variability can be expressed as standardized year-to-year difference or as coefficient of variation (expressed in %). Application of the time-related regression models and their residuals is possible as well. In accessible time series, homogeneity (or non-homogeneity) in a year-to-year variance indicate representativeness and stability of the identified time trends. The time series reported from various sites can be merged for more powerful trend change analysis only if their homogeneity was proved.

7.4 Power analysis

A power analysis is an obligatory step defining magnitude of changes reliably detectable by the statistical methods. Power analysis minimizes risk of the wrong conclusion, misinterpretation or confusing generalization of the observed values. Although its value is widely known, it seldom receives attention in the environmental literature. However, it is essentially vital for establishing program credibility, for supporting effectiveness analyses, and for creating a permanent record of its data for future use.

The relevance of data should be scored mainly with respect to a required decision-making process. In case of global POPs monitoring it is evidently the power of collected data to distinguish substantially increased or decreased POPs levels in the examined matrices and territories in time. Therefore, any time trend analysis must be accompanied with power analysis and the identified trends must be always reported together with corresponding minimum detectable difference. The power analysis estimates a minimum difference between two annually aggregated concentration values detectable by paired t-test on log-transformed data ($\alpha = 0.05$ and $\beta = 0.20$). Appropriate non-parametric alternative like Wilcoxon-rank-sum test or Mann-Kendall test can be used as well, especially when the analysis is based on primary, rather than normalized concentration data.

7.5 Stochastic identification of the time trends

The time trends are identified via qualitative test of statistical significance of the time-related changes observed in the consecutive measurements. At least five consecutive annually aggregated concentration values are required when assessing time trends using one of the following robust techniques:

- The Daniel’s test as an application of the Spearman’s rank correlation coefficient between the concentration values and corresponding time ranks.
- The Mann-Kendal test as a non-parametric test for detection of a trend in the time series based on binary coding of the changes in measurements consecutive in time.

Direction of the time trend (concentration values increasing or decreasing in time) has to be recorded whenever it is confirmed as statistically significant. In addition, any concentration change over time should be reported in the same way, although there is no exact statistical significance behind. Both statistically significant and non-significant time changes over time must be correctly quantified in the reports and marked with p value generated by appropriate tests (see Chapter 7.6).

7.6 Quantification of time trends

Quantification of a time trend should be performed whenever the proper statistical tests confirm significant and consistent time-related differences in the POPs concentrations. Quantified trend means a difference $\Delta=y_1-y_2$, where y_1 respectively y_2 correspond to the annually aggregated concentration values recorded in two different consecutive years. We should be aware of the fact that the recent trend estimates based on just two points („start–end“) differencing can be biased by random errors or fluctuations behind such occasionally measured values. Recent trends should be also related to historical changes occurred in the past because historical trends determine observable changes in more recent periods.

8 Example of standards implemented in global monitoring of atmospheric pollution by POPs: data repository with automated data integration and reporting

The illustrated conceptual model (Table 1) as well as robust statistical processing and reporting of POPs concentration data (Table 2) have already been implemented in a functional data repository and SW toolkit forming IT infrastructure of the Global Monitoring Plan of the Stockholm Convention as defined in Chapter 3 of this paper [22], [24-25]. The methodology, data outcomes and user methodical guide are published in the portal www.pops-gmp.org [31].

The information system distinguishes object entities (nomenclature items) and enables users to stratify accessible measures (content of resources) across a set of classifiers and methodically important attributes. Maintenance of data in an online system is stable and associated with allocated resources for maintaining global connectivity, timely migration to new hardware and inevitable content updates.

The system serves as an example of IT background of global infrastructure for global monitoring and data management. The created data repository is capable of active accumulation of datasets from various sources and ensures their long-term accessibility. Adoption of standards documenting content of the databases, both at the level of data and metadata, enhances usability of the data even for most demanding endpoints, such as long-term trend assessment.

In the centralized data repository, the metadata description is inextricably linked to the datasets, archived and managed in conjunction with the data. The coding of site and data description finally leads to a knowledge-based system which involves gath-

ering and encoding of various types of missing values, values lower than limits of quantification, thesauri, classification rules. Consequently, it allows for implementation of some codes driving automated pre-processing and reporting of imported data which acts upon this coded additional information.

Finally, on-line working interactive communication and reporting level allows users to keep control over the data and to use them, e.g., for regional purposes. The system is equipped with flexible interactive on-line browsing and reporting instruments which are customized for the primary datasets in the data warehouse. Development of such interactive retrieval of data required high degree of familiarity with the data structure and its interpretation; the usage of reports, however, is intuitive and does not limit users, even those without any experience in the field. The outcomes of the reports are documented in Figure 1.

9 Conclusions and future challenges

In this study, we summarized main problems associated with the assessment of performance and impact of long-term environmental programs. Lack of data standards, incompleteness of archived datasets and insufficient statistical power were identified as main limits of functionality of monitoring networks. It is evident that we cannot carry out adequate assessment of the large scale systems with techniques that have been successful in smaller systems with limited heterogeneity [32].

Progress in environmental monitoring, which has been currently advanced by molecular and genomic techniques, need to be supported by development of novel type of IT infrastructures [33-34]. These tools should reflect more universal standards of environmental metadata that go beyond the scope of a particular program or case study [35]. Another inevitable priority is the sustainability of the long-term monitoring programs which does not mean only archiving of gathered datasets. The challenging dimension of sustainability, which is not yet adequately reflected in the literature, includes storing and merging of different datasets in a way maintaining their usage value and information potential [36]. Completed data should remain available and usable after the end of projects. The projects must ensure that the data are captured in a structured way, with necessary metadata description [37]. Nowadays, such a goal is best achieved by utilizing an electronic database coupled with appropriate procedures and data flow facilitated by active participation of data managers and scientists [38].

However, there are still many barriers remaining, especially in global monitoring networks. Data heterogeneity, geographic dispersion of sampling places, funding limitations, incompleteness in time series and other factors often lead to localized, “file-based” solutions [39]. Building centralized data-based archives capable of ensuring retrospective collection of data from these partial repositories represents another priority of current strategies for environmental monitoring. Emphasized attention to the structured collection and correct coding of metadata as well as the primary data must become a key component of either functional approach, not only in eco-informatics [40].

Acknowledgements. This work has been supported by the project TB010MZIP058 “Development of the system for spatial evaluation of the environmental contamination”.

References

1. Mickwitz, P.: A framework for evaluating environmental policy instruments: Context and key concepts. *Evaluation* 9, 415-436 (2003)
2. Crabbé, A., Leroy, P.: The handbook of environmental policy evaluation. Earthscan, London (2008)
3. Goodchild, M.F.: Communicating the results of accuracy assessment: metadata, digital libraries, and assessing fitness for use. In: Mowrer, H.T., Congalton, R.G. (eds). *Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing*. CRC Press, pp. 3-15 (2000)
4. Rönnbäck, B. I., Nordberg, M.L., Olsson, A., Östman, A.: Evaluation of environmental monitoring strategies. *Ambio* 32, 495-501 (2003)
5. Gray, J., Liu, D., Nieto-Santesteban, M., Szalay, A., DeWitt, D., Heber, G.: Scientific Data Management in the Coming Decade. *CTWatch Quarterly* 1, 17-26 (2005)
6. Brunt, J.W., McCartney, P., Baker, K., Stafford, S.G.: The future of eco-informatics in long-term ecological research. *The International Institute of Informatics and Systemics IIIS 7. Proceedings of the 6th World Multi-Conference on Systematics, Cybernetics and Informatics*, 14–18 July 2002, Orlando, FL, USA, pp. 367-372 (2002)
7. Dardari, D., Conti, A., Buratti, C., Verdone, R.: Mathematical evaluation of environmental monitoring estimation error through energy-efficient wireless sensor networks. *IEEE. Trans. Mob. Comput.* 6, 790-802 (2007)
8. Svenfelt, A., Engström, R., Hojer, M.: Use of explorative scenarios in environmental policy-making-Evaluation of policy instruments for management of land, water and the built environment. *Futures* 42, 1166-1175 (2010)
9. Mermet, L., Bille, R., Leroy, M.: Concern-Focused Evaluation for Ambiguous and Conflicting Policies: An Approach From the Environmental Field. *Am. J. Eval.* 31, 180-198 (2010)
10. EPA (Environmental Protection Agency): EPA requirements for quality assurance project plans. EPA QA/R-5, United States Environmental Protection Agency, Washington, DC (2001)
11. EPA (Environmental Protection Agency): STORET, EPA’s largest computerized environmental data system, <http://www.epa.gov/storet/> (2008)
12. Michener, W.K., Brunt, J.W.: *Ecological Data: Design, Management and Processing*. Blackwell Science, Oxford (2000)
13. Jones, M.B., Schildhauer, M., Reichman, O.J., Bowers, S.: The New Bioinformatics: integrating ecological data from the gene to the biosphere. *Ann. Rev. Ecol. Evol. Syst.* 37, 519-544 (2006)
14. Elmagarmid, A., Rusinkiewicz, M., Sheth, A.: *Management of Heterogeneous and Autonomous Database Systems*, vol. 4. Morgan Kaufmann, San Francisco (1999)
15. Grossman, D.A., Frieder, O.: *Information Retrieval: Algorithms and Heuristics*. Springer, Heidelberg (2004)
16. Darwin Core: Darwin Core Schema (version 1.3), a draft standard of the Taxonomic Database Working Group (TDWG), <http://wiki.tdwg.org/DarwinCore>
17. DCM I: DCM I Metadata Terms, <http://www.dublincore.org/documents/dcmi-terms>

18. Huang, P. S., Shih, L. H.: Effective environmental management through environmental knowledge management. *Int. J. Environ. Sci. Tech.* 6, 35-50 (2009)
19. Madin, J., Bowers S., Schildhauer M., Krivov, S., Pennington, D., Villa F.: An ontology for describing and synthesizing ecological observation data. *Int. J. Ecol. Informatics* 2, 279–296 (2007)
20. Williams, R.J., Martinez, N.D., Golbeck, J.: Ontologies for ecoinformatics. *J. Web Semant.* 4, 237-242 (2006)
21. Hale, S.S., Hollister, J.W.: Beyond data management: how ecoinformatics can benefit environmental monitoring programs. *Environ. Monit. Assess.* 150, 227-235 (2009)
22. Stockholm Convention on Persistent Organic Pollutants (POPs): Interim Secretariat for the Stockholm Convention, United Nations Environmental Programme (UNEP) Chemicals: Geneva, Switzerland, Oct 2001, <http://chm.pops.int>
23. Klanova, J., Harner, T.: The challenge of producing reliable results under highly variable conditions and the role of passive air samplers in the Global Monitoring Plan. *Trends Anal. Chem.* 46, 139-149 (2013)
24. Secretariat of the Stockholm Convention: UNEP Report of the First Expert Meeting to update the Guidance on the Global Monitoring Plan for Persistent Organic Pollutants, <http://chm.pops.int/Programmes/GlobalMonitoringPlan/Meetings/GMP1stExpertMeeting2010/tabid/760/ctl/Download/mid/3261/language/en-US/Default.aspx> (2010)
25. United Nations Environment Programme: Guidance on the Global Monitoring Plan for Persistent Organic Pollutants. Secretariat of the Stockholm Convention on Persistent Organic Pollutants, Geneva (2007)
26. Berkley, C., Jones, M.B., Bojilova, J., Higgins, D.: Metacat: a schema-independent XML database system. *Proc. of the 13th Intl. Conf. on Scientific and Statistical Database Management*. IEEE Computer Society (2001)
27. Borgida, A.: Description logics in data management. *IEEE Trans. Knowl. Data Eng.* 7, 671–682 (1995)
28. Cook R.B., Olson R.J., Kanciruk P., Hook L.A: Best practices for preparing ecological data sets to share and archive. *Bulleting of the Ecological Society of America* 82, 138-141 (2001)
29. Cornélis, B., Brunet, S.: A policy-maker point of view on uncertainties in spatial decisions, In: Shi, W., Fisher, P.F. and Goodchild, M.F. (eds.) *Spatial Data Quality*. Taylor & Francis, pp. 168-185 (2002)
30. Hartigan, J. A., Hartigan, P. M.: The Dip Test of Unimodality. *Ann of Stat.* 13, 70-84 (1985)
31. Hůlek, R., Jarkovský, J., Borůvková, J., Kalina, J., Gregor, J., Šebková, K., Schwarz, D., Klánová, J., Dušek, L.: Global Monitoring Plan of the Stockholm Convention on Persistent Organic Pollutants: visualization and on-line analysis of data from the monitoring reports [online]. Masaryk University, <http://www.pops-gmp.org/visualization> (2013)
32. Christensen, S.W., Brandt, C.C., McCracken, M.K. Importance of data management in a long-term biological monitoring program. *Environ. Manage.* 47, 1112-1124 (2011)
33. Emerson, K.J., Merz, C.R., Catchen, J.M., Hohenlohe, P.A., Cresko, W.A., Bradshaw, W.E., Holzapfel, C.M.: Resolving postglacial phylogeography using high-throughput sequencing. *Proc. Natl. Acad. Sci. USA* 107, 16196-16200 (2010)
34. Sork, V.L., Waits, L.: Introduction: contributions of landscape genetics approaches, insights, and future potential. *Mol. Ecol.* 19, 3489-3495 (2010)
35. Metzger, K.J., Klaper, R., Thomas, M.A.: Implications of informatics approaches in ecological research. *Ecological Informatics*, 6, 4-12 (2011)

36. Ozmen-Ertekin, D., Ozbay, K.: Dynamic data maintenance for quality data, quality research. *Int. J. Inform. Manage.* 32, 282-293 (2012)
37. Michener, W.K.: Meta-information concepts for ecological data management. *Ecol. Inform.*, 1, 3-7 (2006)
38. Horsburgh, J.S., Tarboton, D.G., Piasecki, M., Maidment, D.R., Zaslavsky, I., Valentine, D., Whitenack, T.: An integrated system for publishing environmental observations data. *Environ. Modell. Softw.* 24, 879-888 (2009)
39. Bertzky, M., Stoll-Kleemann, S.: Multi-level discrepancies with sharing data on protected areas: What we have and what we need for the global village. *J. Environ. Manage.* 90, 8-24 (2009)
40. Michener, W.K., Jones, M.B.: Ecoinformatics: supporting ecology as a data-intensive science *Trends Ecol. Evol.* 27, 85-93 (2012)

Table 1. Conceptual model proposed for environmental monitoring focused on persistent organic pollutants (POPs) *

CONCEPTUAL MODEL - LEVELS	DEFINITION & COMMENT
RESOURCE IDENTIFIERS	
	Obligatory descriptors identifying institution (project) which guarantees the data (mostly also as owner of the resource content). In already closed resources, the identification is supplied with overall time/ space description.
OBJECTS – KEY ENTITIES	
NOMENCLATURE	Internationally used nomenclature of POP compounds [23-24]. It allows selection of individual compounds and their groups.
OBJECT CLASSIFIERS	Categorized classifiers derived from external (encyclopedic) sources of information. Classifying criterion is linked to the individual compounds and/or to their groups. Classifiers typically define groups but can be used also for scoring of individual records.
Attributes (physical properties, carcinogenicity, etc.)	Attributes extending recognition of nomenclature classes, coded as no/yes/suspected. Code is directly interlinked with individual compounds. The information can be extracted from accessible international databases.
Reference concentration values	Internal (time series analysis, background values) and/or external reference benchmarks. The classifier is coupled with given entity (compound), typically with direct link to matrix sampled and method used.
OBSERVATION – MEASUREMENT (OM)	
TIME & SPACE COORDINATES	Obligatory attributes, also proposed as inescapable items of any data standard.
STUDY TYPE (design)	Study type (list): Long-term environmental monitoring / Short-term environmental monitoring / Case studies / Screening.
PROBLEM STUDIED (exposure)	Problem studied (list): Accident, short-term exposure / Long-term exposure / Random inspection (survey of some area) / Examination of background (reference) site
METHODICAL ATTRIBUTES	Obligatory identification of observation – measurement, necessary for interpretation of measured values. Measured entities select nomenclature items which are examined in given environmental matrix (soil, sediment, water, air, biota). Experimental unit identifies context of measured values (micro-samples within site, site – single sample, site – mixed sample, sample mixed across sites).
Measured entities	Sampling and analytical methods fulfill minimized list of items which follows standardized norms and guidelines.
Matrix	
Experimental unit	
Sampling methods	
Analytical methods	

CONTENT

Measures	Content of the resource, in case of POPs mostly concentration
Units	levels in internationally standardized unit scales. Precision
Precision measures	measures include sample variability (in concentration units) or detection limits of performed analytical methods.

* Example based on the solution adopted in the Global Monitoring Plan of the Stockholm convention [22], [24], [31]

Table 2. Components of robust set of statistical methods recommended for long-term monitoring studies *

COMPONENT		ASSUMPTIONS – OUTCOMES
I. Data and metadata pre-processing		
Assumptions		Recorded primary data in the database (not only e.g. annually aggregated values); accessibility of variability measures for any type of central summary statistics
Outcomes		Initial data filtering according to the objective entities such as site-matrix type and analyzed compounds. The filters also check/verify completeness of the primary database records (sample distribution profile, reported sampling frequency, no. of detected LOQs, LOQs handling rules). Identification and exclusion of proved extreme or unreliable values (e.g. checking their quantile position in the sample distribution function, probabilistic reconstruction of sample distribution, etc.).
II. Summary statistics and baseline estimates		
Assumptions		Correct coding of primary data (direct sampling records) or aggregated values with proper variability measures Baseline estimates of quantitative values (here concentration of POPs for given time period. Two approaches are principally recommended for an assessment of the concentration summary statistics:
Outcomes		a) Median estimate reported together with a 5 th -95 th percentile range b) Geometric mean estimate based on the log-transformed annual averages with a corresponding 95% confidence interval
III. Uncertainty analysis		
Assumptions		Correctly spatially referenced measured values, accessible variability measures for any kind of summarized statistics Information on intra-regional and inter-regional homogeneity Spatial and temporal benchmarking of the sites (or other defined territories)
Outcomes		Check of merging of data from different sites, time periods, etc. Sensitivity analysis tracking changes in information outcome when the database is changed (updated, split, etc.)
IV. Power analysis - effect size estimates		
Assumptions		Evidence – based design of the program: i.e. known sampling strategy, type of collected data, primary end-point and estimate of expected ranges in measured values. Possibility to test the power calculation in a pilot study performed on available primary data sets.
Outcomes		Effect size estimate: quantification of the minimum detectable difference as a base for relevant estimates spatial differences, changes over time and, if possible, the time trends. Defined credibility of a program for future use of its data.
V. Time trend identification		
Assumptions		At least 4-5 consecutive time points (typically years) reported for the examined time series. Possibility to filter out confounding or masking components like seasonality.
Outcomes		Statistically proved time trend and its direction.
VI. Time trend quantification		
Assumptions		Consecutive time points with known variability behind used values (e.g. annually aggregated averages). For relevant estimates, at least 4-5 consecutive time points are preferred rather than two point differencing.
Outcomes		Time-related difference in the measured value, expressed with the following attrib-

utes:

- difference as an absolute value expressed in the concentration units
 - relative (%) change expressed as an index of the value detected in the initial time point (year)
 - 95% confidence interval of the time-related difference
 - p value of the trend test
 - corresponding minimum detectable annual difference
-

* Example customized for time series of POPs concentrations in the Global Monitoring Plan of the Stockholm convention [22, 24, 31]

Fig. 1. Global Monitoring Plan of the Stockholm Convention on Persistent Organic Pollutants: implemented reporting and on-line retrieval of data. Available from: www.pops-gmp.org [31].

