The MESAP Software for the German Emission Inventory

An integrated information system for analysing and reporting statistics and for establishing internet based communication.

Christoph Schlenzig¹

Abstract

Emission reporting is a complex task. Large quantities of data have to be handled. They come from different sources and are communicated through a heterogeneous variety of means and tools. As a consequence they need a high degree of manual treatment which opens up the door for many types of errors. Since most of the reporting is done with Excel, the degree of data redundancy is high, each spreadsheet containing a copy of the original figures. Data flow, data processing and the calculations performed are rarely documented which makes it hard to verify the reported data. The German central system of emissions CSE was designed to achieve the five objectives of emission reporting: transparency, consistency, completeness, comparability and accuracy of the reported data. The requirements were to store time series with different time resolution, to quickly retrieve data and to easily create user defined views on the data. Finally the reporting was to be done using MS Excel format. CSE speeds up data processing and improves the documentation of data and calculation procedures. CSE offers a better management and accessibility of the time series data and allows even to open up this treasure via Internet. The data is well protected and documented in a central database. CSE gives the staff versatile analysis capabilities and lets them create reports in a well known fashion – still with Excel, but with a link to the database. The different steps from collecting new data up to publishing reports are automated as much as possible. CSE is built using the MESAP software technology, which enables the agency to profit from new developments and new functional modules at low cost.

¹ Seven2one Informationssysteme für Energie- und Umweltplanung GmbH
Durlacher Allee 53, D-76131 Karlsruhe
Fon: +49.721.6271400, Fax: +49.721.6271420
christoph.schlenzig@seven2one.de, www.seven2one.de
1. **International Integration**

Germany as a member of the European Union (EU) has to report on air emissions to the European Environmental Agency (EEA) within the framework of the European Regulation 1210/90, which involves the Establishment of an European emission inventory CORINAIR (Core Information of Air Emissions in Europe). The European inventory covers emission data including activity rates and emission factors for 30 pollutants. Specific emissions on cities and districts are required for altogether 435 emission-causing activities.

2. **National Requirements**

The creation of a new database system for the German Emission Inventory was decided by the Federal Environmental Agency in 1997. It was envisaged, that the system should support different data types (e.g. emissions, activity rates or emission factors), which could have different aggregating levels. The system should also offer a high degree of flexibility and expandability.

As a first step the whole data material available at the Federal Environmental Agency (UBA) was analysed and structured into a consistent “matrix” by the Öko-Institut in Berlin. Based on this structure the Institute for Energy Economics and the Rational Use of Energy (IER), University of Stuttgart developed a design for the implementation of the CSE and tested it using a demonstration prototype.

Presently, the CSE software is being implemented based on the MESAP software (Modular Energy Systems Analysis and Planning) developed since 1984 at IER. Seven2one Information Systems, a spin-off company from the University of Stuttgart now in charge of the MESAP development, is responsible for the final implementation of the CSE software.

3. **Organization of Data Flow**

The CSE generates reports of the current emission situation and provides incriminating projections in the necessary degree of detail required by the reports. At present the German Emission Inventory is filled with data by the in-house experts for the main source categories energy, transport, agriculture, production processes etc. Values for emissions, activities and emission factors per year are all stored in the same time series oriented database. Data suppliers are the Federal Statistical Office Germany, research institutes and the owners of industrial plants. The information flow will be organized by a “National System”, an organizational scheme which is being developed at the moment.

The Central System of Emissions CSE consists of an emission reporting database and a database for point source emissions PoSo. The CSE database contains time se-
ries of the main source categories in accordance with the international (UNFCCC, UNECE/EMEP) and national reporting requirements. The PoSo database contains the emissions stemming from power plants and large industries. Figure 1 illustrates the data flow and the modular structure of the German emission inventory.

![Fig. 1. Data flow and structure of the German Emission Inventory](image)

### 4. Software Architecture

CSE is establishing a central data pool on emissions open to all co-workers of the agency. The core of CSE is a powerful, client/server based relational database system for time series oriented data. This data warehouse is highly customisable and stores time series, calculation procedures, the definition of integrity checks and the Excel based reporting formats.

Time series are stored using a customisable multidimensional key. A time series is defined as an indicator reported over time. Each time series has a name and is described by a multi-dimensional time series key. The key is made up of n-dimensions (e.g. attribute, pollutant, spatial unit, sector etc.) describing the meaning of the time series and classifying it according to different criteria and aspects.

Each dimension has its own hierarchical list of specific key words - so called descriptors (e.g. dimension Country, descriptor Germany/Bavaria etc.) which enable to describe each time series as accurate as possible. Because of the hierarchical structure the data can be managed on different aggregational levels.

The number of the dimensions and the entries in the keyword lists can be adapted by the database administrator in order to expand the contents of the database without programming.
A lot of functionality is available around a time series. Time series can “exist” at any time resolution (yearly, monthly, daily, hourly, etc.). Aggregation rules convert time series to higher levels (from monthly values to yearly aggregate). Every time series can have its own “mapping” rules to interpolate missing values. Hypothesis can be defined to store different projections of the time series for the future. The values of the time series can be converted in any other compatible physical unit for display, reporting or calculation using the SI system (Système International). Time series specific plausibility checks can be defined for data entry.

Each data value within a time series can have its own documentation, which is structured in different components (e.g. comments, quality key, editor, literature references etc.). The documentation can be configured by the database administrator without programming.

Flexible retrieval functions based on this multidimensional key allow the user to create his own views on the data. Based on these views, CSE offers a set of user friendly tools to facilitate data entry and to allow for extensive documentation of the origin and quality of the data. A set of interfaces allows to link to other existing electronic data processing systems.

The CSE separates data management from calculation procedures and reporting. Calculation procedures and reports contain only variables linked to the corresponding time series in the database. Only when the calculation is executed or the report is generated, the actual data from the database is inserted. This architecture avoids the problem of data redundancy when storing data in Excel-Sheets. Report generation is still Excel based, but with hotlinks between Excel and the central database. A special add-in function makes it easy to shift the reporting period for a whole sheet. This allows a high degree of flexibility for creating user defined reports containing tabular evaluations and additional calculations for periodical publications. Thus CSE can be easily adapted to new reporting requirements without modification of the underlying database structure and without programming.

5. The MESAP modules

The DataSheet is the main tool for data entry and retrieval (figure 2). The DataSheet can be used to create user definable views on the data by setting filters for the different dimensions. Since the keywords in every dimension are arranged in a hierarchical list similar to a tree structure, filters can be set by selecting a node in this tree. The DataSheet displays all time series corresponding to the search pattern established by the filter. The hierarchies allow to search for groups of keywords. Finally it is also possible to perform a simple text search. At the centre one can see the multidimensional key describing the time series. On the right side one can see the grid where data can be entered. On the bottom is a graphics window which allows to visualize selected time series. All filter and layout settings defined by a user
Fig. 2. The DataSheet as the main navigation tool of the CSE

can be stored as a so called view. The window on the left (inspector panel) shows a list of available views. An internet based DataSheet is also available to provide Internet access to the CSE data.

The Analyst is the report generator of the CSE to fulfil the reporting requirements. The Analyst is Excel based and thus allows the user to create his report templates in a well known procedure. The major difference is, that the reported values are not copied into the sheet but are inserted by establishing a hot link to the database. This allows to use all advantages of Excel together with the benefit of a centralized and documented data management. Reports can be updated by the push of a button with the most recent figures from the database. In addition, the time shift functionality allows to change the whole reporting period. Creating the report for the next year becomes a matter of pushing a button – after the data has been entered into the database. All report templates are stored in the CSE database.
The **CalQlator** allows to specify calculation procedures which estimate the emissions from other indicators by some mathematical correlation. In a first step the user establishes the mathematical equation system and specifies which input variables for the calculation correspond to which time series in the database and which time series are to be used to store the results. This is done by using the equation editor. Once the calculation procedure is defined correctly and stored in the database, it can be executed by the CSE solver. Several calculation procedures can be grouped together in a hierarchical structure. This way complex calculations can be broken up transparently into smaller blocks that are easy to control. A status system indicates, whether the results of a calculation are up-to-date or whether the calculation has to be re-launched. The execution of calculations can also be scheduled. The results of the calculation can be incorporated in reports by establishing a link to the resulting time series of the calculation procedure.

CSE provides a high degree of flexibility concerning the mathematical formulation of the calculation procedures. The calculation procedures are stored and documented separately in the database. Integrity checks are available during data entry as well as for testing the calculated results.

MESAP is a client/server based software running on Windows NT/2000/XP. It uses an object oriented 3-tier architecture and is .NET compatible. Three DBMS platforms are supported: MS Access 2000, MS SQL Server 7 and Oracle 8i. MESAP offers an API for the development of customized solutions. MESAP is multi user capable and offers a detailed access right management.