

Spatio-temporal analysis of environmental pollution in urban areas: A case study of the environment in the city of Prague

L. Matejicek^a

^a *Institute for environmental studies, Charles University in Prague*
Email: lubos.matejicek@natur.cuni.cz

Abstract: Environmental protection of urban areas requires a wide range of data about air pollution, water pollution, soil pollution, waste management, noise, transport, landscape, and energy supply. In order to perform impact assessment and model predictions, various techniques of short-term, medium-term and long-term monitoring are required for spatial processing and time series analysis. The Geographic Information System (GIS), with data from remote sensing, monitoring networks and field sampling, is used to create datasets. In order to raise GIS spatial processing and dynamic modelling to a more efficient level, new data management tools are developed for basic data management, more complex exploratory spatio-temporal data analysis (ESTDA), dynamic modelling by partial differential equations and ordinary differential equations, and cell-based modelling focused on multicriteria analysis. In order to accelerate numerical tasks, the data structures are based on information depicted in 3-dimensional space, in a data cube, where two dimensions correspond to the surface coordinates, and the third corresponds to time. One of the strongest aspects of environmental information arranged into a data cube is its capability to support analytical methods derived from spatial modelling and dynamic modelling. The data cube is made up of cells. Each cell occupies a specific portion of 3-dimensional space in a regular grid. It can represent spatio-temporal data predicting the values of air pollution concentration, or data predicting another type of environmental pollution such as surface water pollution, soil contamination, noise, or other disruptions of our environment. Multicriteria analysis with cell-based modelling helps to find the worst sites characterized by high levels of environmental pollution. This can be derived from interpolations based on sample points of terrain measurements. The methods must reclassify the datasets to a common scale, and weight those that are more important for the analysis. Finally, the dataset represented by the weighted average of individual factors is used to identify the most exposed sites. In fact, Inverse Distance Weighted (IDW) is implemented as a deterministic method, and ordinary kriging is used for geostatistical interpolation. While the chosen interpolations in 3-dimensional space are based on relatively simple numerical tasks and statistical principles, information in cells should be supported by more powerful modelling tools that can encompass transport (consisting of advection, diffusion and dispersion), sorption, decay or degradation, reaction (either kinetic or thermodynamic) and other phenomena. In case of air pollution spreading in the atmosphere, the effect of diffusion, advection in a constant unidirectional flow field, and decay are described by an analytical solution derived from partial differential equations. Mathematical models based on partial differential equations are often transferred to the simplified form of a set of ordinary differential equations. Surface water pollution is modeled with this simplified form that deals with mass accumulation and delayed transport with compartment models. The attached case study describes the environment in the City of Prague and characterizes the basic monitored compartments of environmental pollution, such as air pollution, surface water pollution and land-use changes. A few case studies are described to show data processing and spatio-temporal analysis of short-term, middle-term and long-term observations. A short term spatio-temporal analysis of air pollution in the central part of the City of Prague is focused on producing interpolated continuous surface maps from point samples and from air pollution models during a single day. A medium-term evaluation of surface water pollution illustrates changes in N-NO₃ concentrations in selected profiles over the period of a decade based on compartment models. A long-term spatio-temporal analysis describes land-use changes derived from classified aerial images in the area of interest in the 1938-2006 period. The presented spatio-temporal analyses indicate changes in living conditions in the city of Prague, which can be useful in decision-making processes and outline common guidelines for environmental predictions. It also demonstrates spatio-temporal tasks in dependence on various time scales and spatial scales.

Keywords: *Spatio-temporal analysis, interpolation, dynamic modelling, environmental pollution*

1. INTRODUCTION

Spatio-temporal analysis and modelling address problems that are spatially distributed (Zeiler, 2010) as well as temporally related (Jørgensen and Bendoricchio, 2001). The two basic methodologies addressing these dimensions are spatial databases managed by the Geographic Information System (GIS) and dynamic models processed by simulation tools on the other hand (Maquire et al., 2005). An additional dimension is added by integration with the monitoring data (Fedra, 1999). Most urban environmental considerations dealing with environmental problems such as air pollution, surface water pollution, soil pollution, noise, waste management and the landscape (Kamp et al., 2003) have a spatial dimension that can be addressed by GIS (Mitchell, 2005). In addition to spatial management and analysis, current GISs can also support high speed numerical calculation in order to solve dynamic models (Goodchild et al., 1996). Thus, the spatial dimension focused on both geometric data (coordinates and topological information about points, lines and polygons in thematic map layers) and attribute data (information describing the properties of spatial objects) can be extended by data structures and modules for spatio-temporal analysis and modelling (Matejicek, 2010).

2. MULTICRITERIA ANALYSIS AND MODELLING IN SPACE AND TIME

Spatio-temporal data can be stored in a variety of formats. The spatially distributed data are stored in databases that offer advanced data structures and efficient management tools. The case-oriented geo-database is the native data structure for the storage and analysis of geographic information. The geographic features are incorporated in the thematic layers that represent real-world entities. The thematic layers can also include datasets representing satellite and aerial photo imagery. They are used as a foundation for other thematic layers or to represent continuous surfaces. The geographic features and imagery are georeferenced so that they can also be aligned (Maidment, 2002).

Relational database management systems (RDBMSs) support complex data structures, such as GIS datasets and their topology. In time-consuming numerical calculation, direct RDBMS support slows down simulation of dynamic models. In order to accelerate the data storage, a variety of binary formats is used for direct data management in simulation tools (Holzbecher, 2007). The next step can be represented by parallel computing tools that utilize parallel memory buffers for high speed computations. Various implementations of parallel calculations on personal computers are based on multithread computing in central processing units or graphic processing units adjusted for simulation tasks (Kirk and Hwu, 2010).

In spatio-temporal analysis of environmental pollution in urban areas, the data are arranged in cubes in order to allow fast multicriteria analysis and modelling in space and time. Although it uses similar principles to RDBMSs for online analytical processing (OLAP) (Liang, 2000), structured query language (SQL) is not practical due to the high speed requirements of dynamic modelling tools. The basic structure is based on the information depicted in 3-dimensional space, where two dimensions correspond to the surface coordinates (indexes x and y), and the third corresponds to time (index t). Figure 1 shows an example of environmental information about surface air pollution depicted in 3-dimensional space. Generally, a number of other dimensions can be adjusted to coordinates such as the 3rd coordinate of space (index z) or dimensions derived from various spatial scales and time scales (index s). The 3-dimensional data cube is made up of cells. Each cell represents a specific portion of 3-dimensional space in a regular grid. All the cells are the same size and small enough for performance of the most detailed analysis. A number of steps must be carried out to create a data cube. After setting the data extent and resolution, the input data, such as measurements of environmental pollution and simulation outputs from dynamic models, have to be converted to attributes of discrete data represented by a set of cells in a data cube grid.

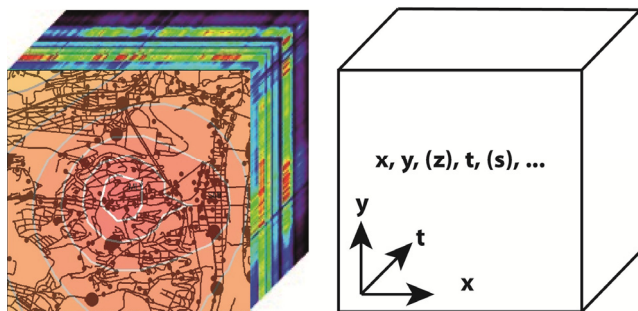


Figure 1. An example of environmental information about surface air pollution depicted in 3-dimensional space, where two dimensions correspond to surface coordinates (indexes x and y), and the third corresponds to time (index t) in the 3-dimensional data cube.

One of the strongest aspects of environmental information arranged into a data cube is its capability to support analytical methods derived from cell-based modelling. The value associated with each cell identifies the type of phenomenon that is relevant at each location. It can represent spatio-temporal data predicting the values of air pollution concentration based on the sample points and data series, or data predicting another type of environmental pollution, such as surface water pollution, soil contamination, noise, or other disruptions of our environment.

The initial step of analytical research includes Exploratory Spatio-Temporal Data Analysis (ESTDA), which enables deeper understanding of the phenomena we are investigating. ESTDA is composed of a series of tools allowing a view into the spatio-temporal data. It can explore the distribution of the data and look for global and local trends. Orientation among multiple datasets in the data cube can be supported by multiple regression and correlation. A number of other techniques can be implemented for identification of time series by curve-matching methods in order to explore seasonal changes in environmental pollution.

Multicriteria analysis using cell-based modelling helps to find the worst possible locations characterized by high levels of environmental pollution. It is derived from interpolations based on sample points of terrain measurements. The methods must reclassify datasets to a common scale, and weight those that are more important for the analysis. Finally, the dataset represented by the weighted average of the individual factors is used to identify the most exposed sites. As an example in Figure 2, location of potential sites for remediation is derived from 3 datasets of spatio-temporal data. Each dataset is represented by the values of the concentrations of a selected pollutant, such as NO_x, PM₁₀ and SO₂. Each data cube includes spatial interpolations of the selected air pollutant in a chronological sequence. After reclassification of the concentrations in the data cube to a common scale, for example 1-100 where higher values are assigned to less suitable values, the weighted values are combined to find highly exposed sites. The results in a chronological sequence enable us to explore the length and variation of exposure.

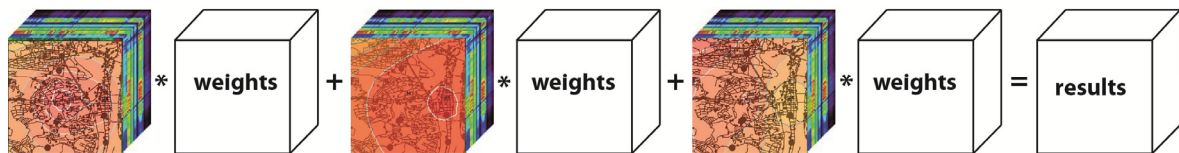


Figure 2. Environmental information (about surface air pollutants such as NO_x, PM₁₀ and SO₂) depicted in the data cubes, where two dimensions are correspond to the surface coordinates (index by x and y), and the third corresponds to time (index by t).

Two main groups of interpolation techniques, such as deterministic and geostatistical, are needed to provide comparable data cubes with the same cell size and grid extension, based on the original attribute data from sample points. In fact, Inverse Distance Weighting (IDW) is implemented as a deterministic method, and ordinary kriging is used as a geostatistical interpolation. Generally both methods are implemented to provide interpolation in n -dimensional space (Matejicek, 2006). In order to demonstrate interpolations in correspondence with the described cell-based analysis, 3-dimensional space (x, y, t) is used for the described tasks.

While the chosen interpolations in n -dimensional space are based on relatively simple numerical tasks and statistical principles in order to complement any data in a data cube, environmental predictions should be supported by modelling tools that can encompass transport (consisting of advection, diffusion and dispersion), sorption, decay or degradation, reaction (either kinetic or thermodynamic) and other phenomena. For air pollution spreading in the atmosphere, the effects of diffusion, advection in a constant unidirectional flow field, and decay are described by analytical solution (Holzbecher, 2007):

$$c(x, y, z, t) = \frac{M}{(\sqrt{4\pi t})^3 \sqrt{D_x D_y D_z}} \exp \left\{ -\frac{1}{4t} \left[\frac{(x - vt)^2}{D_x} + \frac{y^2}{D_y} + \frac{z^2}{D_z} \right] - \lambda t \right\} \quad (1)$$

where M denotes the total mass per unit area in the fluid system, diffusivities D in x, y and z directions depend on the transported pollutant and its environment, velocity v corresponds to constant advection in the x -direction, λ refers to the decay coefficient, and the concentration distribution $c(x, y, z, t)$ around the 3D instantaneous source depends on coordinates x, y, z and time t . The simplified formula is derived from the partial differential equation:

$$\frac{\partial c}{\partial t} = \frac{\partial c}{\partial x} D_x \frac{\partial c}{\partial x} + \frac{\partial c}{\partial y} D_y \frac{\partial c}{\partial y} + \frac{\partial c}{\partial z} D_z \frac{\partial c}{\partial z} - v \frac{\partial c}{\partial x} - \lambda c \quad (2)$$

and must be adapted to real urban environmental conditions that include a number of other sources of pollution and specific surface interactions caused by built-up areas. In case of two dimensions with surface coordinates (x, y) and the third formed by time (t) , only the data from the surface layer are exported to the data cube in a chronological sequence. There are many applications of the partial differential equation (2), mainly for dispersion of pollution in the atmosphere. Mathematical models based on partial differential equations are often transferred to simplified forms of compartment models. Surface water pollution can be based on a set of ordinary differential equations that include the mass accumulation and delayed transport with compartment models (Matějček, 2003):

$$\frac{d(cv)_{(i)t}}{dt} = f_{(i-\tau)t-1} + \sum_{k=1}^{ni} f_{(i)ki} - (cv)_{(i)t} \frac{q_{(i)t}}{v_{(i)t}} + f_{(i)Pi} + f_{(i)NPi} - f_{(i)Ri} \quad (3)$$

where variables $f_{(i)t}$ are flows that represent transport of the released pollutant at the present time value and variables $f_{(i-\tau)t}$ are flows that depend on the solution at previous times with the lags τ . The accumulation in the i^{th} compartment depends on inflows and outflows. The most important inflows for the i^{th} compartment are represented by a main upper stream inflow $f_{(i-\tau)t-1}$, subsidiary inflows $f_{(i)ki}$, inflows caused by point sources $f_{(i)Pi}$ and non-point sources $f_{(i)NPi}$. The most important outflow is the downstream outflow $(cv)_{(i)t}q_{(i)t}/v_{(i)t}$, where variable $cv_{(i)t}$ is the amount of transported pollutant (the concentration c_i multiplied by the water volume v_i in the i^{th} compartment), $q_{(i)t}$ is the water flow and $v_{(i)t}$ is the water volume in the i^{th} compartment. The outflow $f_{(i)Ri}$ is caused by chemical and biological interactions that significantly depend on the pollutant and its environment in the i^{th} compartment. Other types of environmental pollution caused by waste management, noise and landscape disruptions require specific forms of descriptions and their predictions are highly dependent on terrain measurements and local decision-making processes.

3. A CASE STUDY OF THE ENVIRONMENT IN THE CITY OF PRAGUE

The City of Prague is the largest city in the Czech Republic with an area of 496 km² and a population of 1.25 million inhabitants in 2009, Figure 3. The factors affecting the quality of the environment are especially the impacts of automobile traffic, water and energy management, waste management, and sustainable land use. The former regime prior to 1989 managed the urban environment in a manner that greatly increased the burden (Prague Environment, 2009).

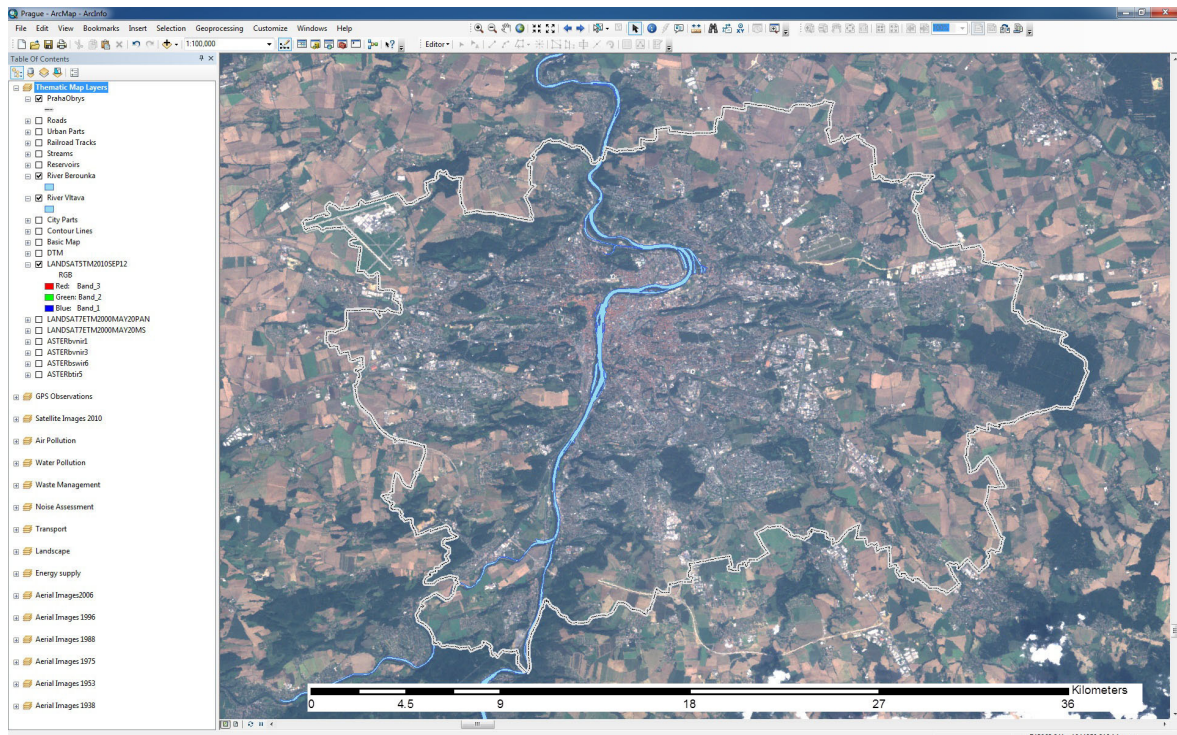


Figure 3. A GIS-based application focused on the environmental information system of the Capital City of Prague on a scale of 1:100,000 with selected map layers: the boundary line of the City, the river network represented by the Vltava and Berounka Rivers and the LANDSAT 5 TM image in the background.

A few analyses were performed to provide challenges for multicriteria analysis to define the parameters of spatio-temporal analysis of environmental pollution in urban areas of the City of Prague. The approaches were based on various spatial scales and time scales in dependence on the examined processes and the available data.

Air pollution - a short-term spatio-temporal analysis: Mapping of human exposures to air pollution from traffic has become an important issue over the past decade. The monitoring network of measuring stations in the territory of the City of Prague archives time series of concentrations of atmospheric pollutants, such as sulphur dioxide, nitrogen oxides, tropospheric ozone, carbon monoxide, suspended particulate matter (PM₁₀) and aromatic hydrocarbons, specifically benzene and its derivatives. The spatio-temporal analysis was focused on processing prediction maps from pollutant point measurements taken at one hour intervals. Selected outputs of the spatial interpolations are depicted in Figure 4.

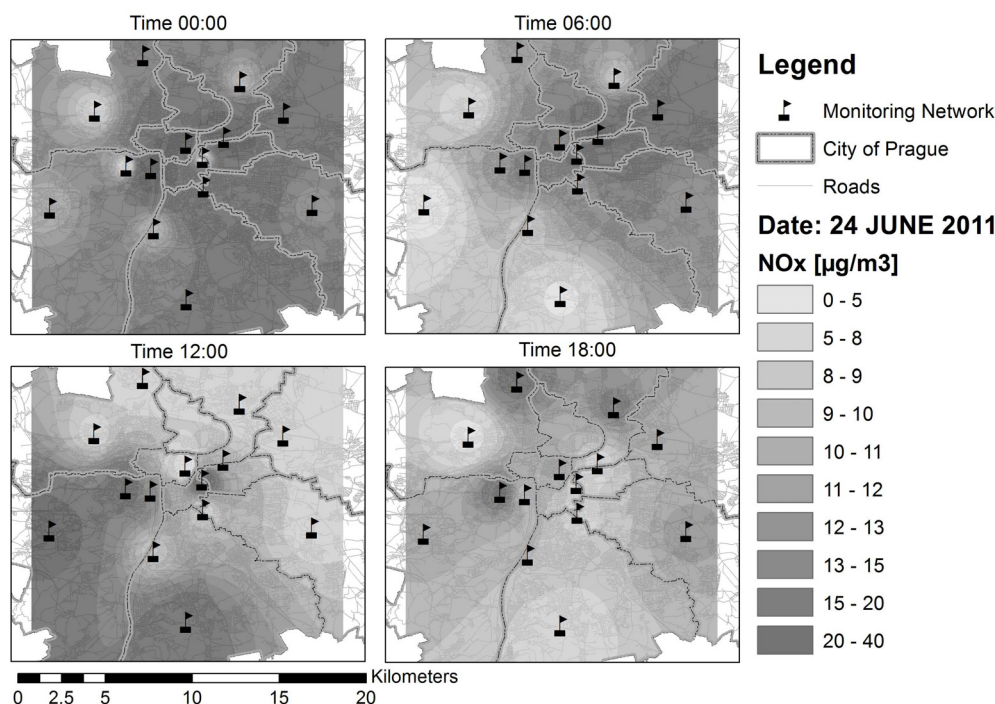


Figure 4. A short-term spatio-temporal analysis of air pollution in the centre of the City of Prague during one day (24-June-2011) producing continuous surface maps from 14 point samples, predicting the values of the NO_x concentration [µg/m³] for every location in the area of interest.

Surface water pollution - a medium-term spatio-temporal analysis: Trends in general, physical and chemical indicators of surface water quality in the city territory are based on regular measurements in the selected profiles of rivers and smaller water courses. Initial data outputs of water quality are processed by the Czech Hydrometeorological Institute. As an example, selected map layers complemented by charts are given in Figure 5, which illustrates seasonal changes in the N-NO₃ concentration [mg/l] at the selected stream mouth taken in a one month interval during the 2001-2009 period.

Land use changes - a long-term spatio-temporal analysis: Similarly as rapid land-use changes take place in many mega cities (Wu et al., 2006), the City of Prague has been undergoing the phenomenon of urban development at the expense of agricultural land over past few decades. Land-use changes in the 1938-2006 period are illustrated in Figure 6. The area of interest is represented by raster datasets derived from aerial images taken in 1938, 1953, 1975, 1988, 1996, and 2006.

4. DISCUSSION AND CONCLUSIONS

Multicriteria analysis and modelling are tested on a number of data sources. Short term spatio-temporal data focused on air pollution enable implementing of numerical models and terrain measurements in a chronological sequence of the surface layers. Model simplification for a few predominant sources of pollution yielded approximate predictions. The spatial interpolations based on sample points yield approximate results outside of the monitored sites. Thus, multicriteria analysis can reveal basic trends in air pollution.

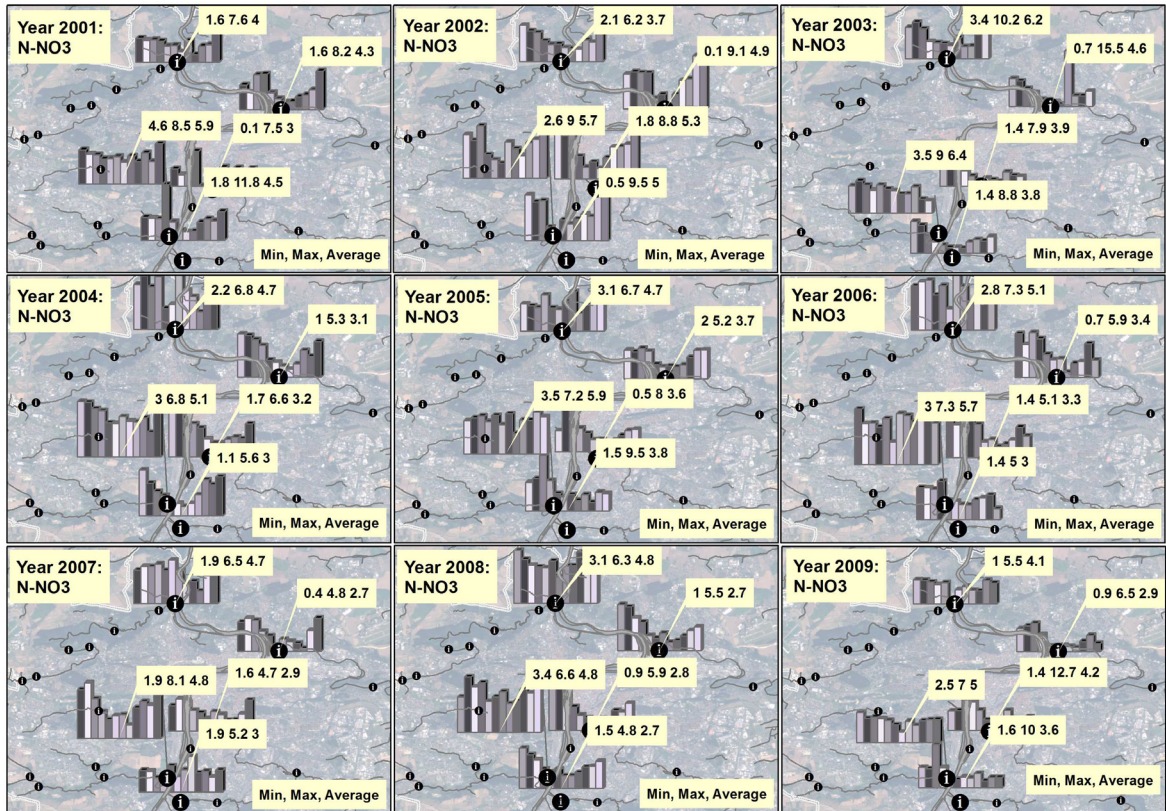


Figure 5. Selected map layers complemented by charts that illustrate seasonal changes in the N-NO₃ concentration [mg/l] at the selected stream mouth taken in a one month interval during the 2001-2009 period.



Figure 6. The land use changes in the area of interest in the 1938-2006 period were derived from aerial images.

A medium-term spatio-temporal analysis of surface water pollution explores changes in pollutants in compartments that represent parts of the surface water network. Multicriteria analysis and cell-based modelling are focused on only some of the cells in the data cube that cover the surface water network. The ESTDA can help to identify correlations among the compartments.

The long term spatio-temporal analysis focused on land use changes in the 1938-2006 period was derived from aerial images dealing with development of the city over a period of a few decades. It is based on classification of the surface into land-use classes. Multicriteria analysis yields identification of trends in the area of interest.

The described methods and the attached case study demonstrate the use of tools for spatio-temporal analysis and modelling using data that are spatially distributed as well as temporally related. A 3-dimensional data cube consisting of cells is used for basic data management, more complex exploratory analysis designated as ESTDA and cell-based modelling with multicriteria analysis. A case study of the environment in the City of Prague is used to demonstrate the described tools using data from model simulations and terrain measurements of environmental pollution in short-term, medium-term and long-term periods. The outlined research can support decision-making processes in the framework of the City of Prague and indicate potential solutions for similar areas of interest.

ACKNOWLEDGMENTS

The author wishes to thank the authorities of the City of Prague, and the Institute of Municipal Informatics for providing the data of the environment in the framework of student case studies. GIS processing was supported in part by the Ministry of Education, Youth and Sports under Grant FRVS 131/2009A/b.

REFERENCES

- Ekins, P. and Lees E. (2008). The impact of EU policies on energy use in and the evolution of the UK built environment. *Energy Policy*, 36, 4580-4583.
- Fedra, K. (1999). Urban environmental management: monitoring, GIS, and modeling. *Computers, Environment and Urban Systems*, 23, 443-457.
- Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C. Maidment, D., Crane M. and Glendinning, S. (1996). GIS and Environmental Modeling: Progress and Research Issues. Fort Collins: GIS World Books.
- Holzbecher, E. (2007). Environmental Modeling Using MATLAB. Berlin: Springer.
- Jørgensen, S.E. and Bendoricchio, G. (2001). Fundamentals of Ecological Modelling. Amsterdam: Elsevier.
- Kamp, I., Leidelmeijer, K., Marsman, G. and Hollander, A. (2003). Urban environmental quality and human well-being Towards a conceptual framework and demarcation. *Landscape and Urban Planning*, 65, 5-18.
- Kirk, D.B. and Hwu, W.W. (2010). Programming Massively Parallel Processors. Amsterdam: Elsevier.
- Liang, W., Wang, H. and Orlowska, M.E. (2000). Range queries in dynamic OLAP data cubes. *Data & Knowledge Engineering*, 34, 21-38.
- Maidment, D.R. (2002). ArcHydro GIS for Water Resources. Redlands: ESRI Press.
- Maquire, D.J., Batty, M. and Goodchild, M.F. (2005). GIS, Spatial Analysis and Modeling. Redlands: ESRI Press.
- Matějčiček, L., Benesova, L. and Tonika J (2003). Ecological modelling of nitrate pollution in small river basins by spreadsheets and GIS. *Ecological Modelling*, 170, 245-263.
- Matějčiček, L. (2006). A GIS-based approach to spatio-temporal analysis of environmental pollution in urban areas: A case study of Prague's environment extended by LIDAR data. *Ecological Modelling*, 199, 261-277.
- Matejicek, L. (2010). Environmental Modeling with GPS. New York: Nova Science Publishers.
- Mitchell, A. (2005). GIS Analysis-Volume2: Spatial Measurements & Statistics. Redlands: ESRI Press.
- Prague Environment (2011). http://envis.praha-mesto.cz/rocnky/Pr09_html/index_an.htm [12-June-2011].
- Schwarz, N. (2010). Urban form revisited-Selecting indicators for characterizing European cities. *Landscape and Urban Planning*, 96, 29-47.
- Starkl, M., Brunner, N., Flögl, W. and Wimmer, J. (2009). Design of an institutional decision-making process: The case of urban water management. *Journal of Environmental Management*, 90, 1030-1042.
- Wu, Q., Li, H., Wang, R., Paulussen, J., He, Y., Wang, M., Wang, B. and Wang, Z. (2006). Monitoring and predicting land use change in Beijing using remote sensing and GIS. *Landscape and Urban Planning*, 78, 322-333.
- Zabel, T., Milne, I. and McKay, G. (2001). Approaches adopted by the European Union and selected Member States for the control of urban pollution. *Urban Water*, 3, 25-32.
- Zeiler, M. (2010). Modeling our World. Redlands: ESRI Press, 2010.