Chapter 8: Modelling of environmental pollution in urban areas with GIS

LIBUŠE BENEŠOVÁ, LUBOŠ MATĚJÍČEK

1. Introduction

Description of principles focused on environmental pollution in urban areas can be developed in a more sophisticated way by dynamic models, which can manage spatio-temporal data. One of the information systems that offer management of data with spatial and time dependent attributes represents Geographic Information System (GIS). Whereas the GIS manage spatial description of urban environment, modelling is focused on functioning of environmental processes, which can include various types of urban pollution. In spite of the dynamic modelling and GIS are synergistic (GIS can serve as data analysis framework for models of pollution), both systems grew up separately. The roots of dynamic modelling come mostly from physics, which offer wide range of deterministic interactions described by physical laws and rules. GIS originate from cartographic endeavour to automate map development and updates. But nowadays, both these individual research areas are linked together with their possibilities to support description of spatial interaction and time phenomena in the frame of environmental models. For all that, the computer programs, which represent dynamic modelling and GIS, have different data structures and functions for data management and numerical calculations. This makes it difficult to link inputs and outputs together to carry out complex analysis.

Environmental protection of urban areas, which includes a wide variety of techniques in order to study individual factors of pollution and stress, overtax many of present data structure formats. New concepts of data links and sharing of functions are needed. In the past, air pollution, contamination of soils and water, noise assessment, waste management and landscape protection were monitored at the local urban scale. Independent computer tools were used to analyse data measurements and to predict effects. At present, environmental complex analysis of urban areas dictates integration of individual dynamic models and data structures to one system. A number of applications is described in the book edited by Goodchild (1996).

General description of dynamic modelling, which is applied in environmental science, is a subject of books published by Jørgensen (1985, 1994), Hannon (1994), Legendre (1998) and Straškraba (1985). The various aspects of extensions of GIS into environmental research describe Burrough (1998) and Johnston (1998). The actual case studies dedicated to environmental modelling and GIS in the urban areas are
presented in growing scale in international journals and meetings (Matějíček, 2002, 2001a, 2000a, 1996a).

2. Environmental pollution in the urban systems: specification of the model components

Environmental modelling contains wide range of experimental research activities. The main aims represent estimates of short term and long term changes in various scales of urban areas, data analysis and development of models that can simulate real situation to aid decision making processes. Due to range of relatively different research methods, the studies of environmental pollution in urban areas are divided into individual disciplines. Modelling of air pollution includes localization of global and local sources, description of particle transport and chemical interaction in atmosphere. Development of surface water and groundwater models together with soil pollution is focused on similar phenomena but in the different physical and chemical environment. In the frame of ecological systems, the air, water and soil pollution is mostly considered as combination of accumulation and transfer phenomena in abiotic parts of ecological systems. Biotic parts can be affected consecutively. The description of spreading pollutant in dynamic models has to be subordinated to ecological interactions. The complex study of urban environment should complement landscape protection, waste management, noise assessment and other phenomena, which can affect living environment in urban areas (biomonitoring, allergy assessment).

2.1 Air pollution

Monitoring of air pollution in larger urban areas belongs to standard routines of environmental assessment. The systems are usually divided on registering of sources of air pollution (emissions), assessment of air quality on the basis of monitoring (immissions) and other relevant information (smog regulation system, meteorological measurements and forecast). The sources of atmospheric pollution are registered into four categories in the REZZO Register. The first three classes comprise stationary sources. The fourth class contains mobile sources. The assessment of air quality is based on the air quality limits for pollutants, which are related to average annual, daily, 8-hour’s and half-hour pollutant concentration. The assessment of air pollution situation is based on data stored in central databases. Mostly monitored pollutants represent sulphur dioxide, nitrogen oxides, ground-level ozone, carbon monoxide, suspended particulate matter and aromatic hydrocarbons. Other information about air pollution can give atmospheric deposits, dust fallout, heavy metals and organic compounds.

A wide range of data and measurements from air pollution monitoring networks has to be stored and effectively processed on computer information systems. Evaluation of actual situation and long term trend of air pollution require integration of wide range of data usually in the form of time series together with geographical informa-
tion about the location of pollution sources and monitoring sites, layouts of residential, industrial and traffic areas. The primary data have to be processed and analysed to derive specific characteristic for decision-making processes. The basic level covers geostatistical procedures (spatial correlation and regression, factor analysis, interpolation). In the frame of the GIS, the map algebra can be applied to estimate more complex spatio-temporal characteristics (risk assessment). The superstructure represents dynamic models that are focused on transport of air pollution and possibly chemical interaction in the urban environment.

Whereas the GIS are used for data management and geostatistical analysis, the dynamic modelling of air pollutants is mostly solved in standalone computer programs, which include implementations of numerical algorithms for solution of partial differential equations (Matějíček, 2000). In spite of wide range of functions, which are implemented in the GIS for spatial data management, existing GIS do not offer effective programming environment for implementation of time and capacity consuming calculations in the frame of dynamic modelling. On the opposite site the GIS offer far more functions for storage and visualisation of data. So the GIS are mostly used in the frame of dynamic modelling of air pollution for pre-processing (data input, data analysis, estimation of model parameters) and post-processing (data management and visualisation, printing maps, presentation of results on the Internet).

2.2 Water pollution

Aquatic ecosystems perform numerous valuable environmental functions. They cycle nutrients, purify water, attenuate floods, recharge ground water and provide habitat for wildlife and recreation for people. Increases population specially in some part of the world, accompanied by intensified industrial, commercial, and resident development have led to pollution of water by fertilizers, insecticides, oil, toxic landfill leachates and feedlot waste.

Water pollution we can divide in characteristic groups according to:
1. type of pollution (organic, heavy metals, thermal pollution, radioactivity)
2. the author of pollution (sewage from people, agricultural, industry)

For our data processing the better is to use categorization according to type of pollution:

A. Organic pollution
The term of organic pollution has two meanings in the environmental context. In the strict chemical sense it refers to chemical compounds of carbon. Organic chemicals are produced from living organism and are also manufactured from coal and oil [both from which were once living].

The second meaning of organic refers to a method of agriculture, which does not involve the use of artificial fertilizers or pesticides.

Organic pollution occurs when large quantities of organic compounds are released into watercourse. During the decomposition process the dissolved oxygen is reduced and it cause the oxygen depletion, which have severe consequences for the stream
biota. Ammonia is often present and this adds to toxicity. Organic waste from man and animals may also be rich in disease-causing pathogenic organism.

The origin of these compounds is: domestic sewage, urban run-off, industrial effluents and farm wastes.

Sewage effluents are the greatest source of organic materials discharged to the freshwater. In urban areas the run-off form houses, factories, and road can resulted in severe pollution. This urban run-off may be routed through sewage works or be combined with sewage system. Where the urban run-off drains directly to the river pollution may be severe, because drainage from the hard surfaces is so rapid that the dilution will be minimal. The quality of run-off is highly variable, but BODs as 7,700 mg/l has been recorded.

Industrial effluents are a further source of organic pollution. These may be routed via waste water treatment plants or without treatment directly into the river. Amongst the industries which produced effluents containing organic matters are first of all food processing and brewing industry, dairies, abattoirs and tanneries, and textile and paper making factories.

Farm effluents have become an increasing pollution problems. It relates with the intensification of livestock production in recent years. The more intensive and mechanized ways of dealing with vegetable crops may also cause pollution.

B. Heavy metals

Heavy metals are a group of elements which have a number of chemical properties in common but whose biological roles range from the essential to the highly toxic. On the one hand copper and zinc are the essential trace elements required for the functioning the body. On the other hand mercury, cadmium and lead have no biological role in the healthy organism and can cause damage even at very low levels. There is, however considerable evidence that some heavy metals may be harmful to health at levels which are found in the environment. They are conservative pollutants in that they are not broken down, or are broken down over such a long time scale, that they effectively become permanent additions to the aquatic systems. They accumulate in organisms. The major uptake route for aquatic organisms is directly from the water. Carnivores at the top of food chain, however, such as birds, mammals, including humans, obtain most of their pollutant burden from aquatic ecosystem by way of their food [especially fish].

C. Thermal pollution

Most electricity-generating plants [using fossil or nuclear fuels] operate through the thermodynamic process, in which high pressure steam is produced and than expanded by turbines. Turbines convert thermal energy into mechanical energy. The condensers require large amount of cooling water, which is removed from ground and surface waters and returned at a higher temperature. Most of this water can be reused and most of it is free of contamination other than the heat. But the increase of temperature is very serious problem, because it alters the physical environment in terms of both a reduction in the density of water and its oxygen concentration,
which varies inversely to with temperature. Except the physical quality the biota is influenced. There are various possible effects of heated effluents on the biology of the waters that receive them. Those species intolerant of warm conditions may disappear, while other species, rare in unheated water, may thrive so that structure of the community changes. Bleu-green algae, for example, are most tolerant and become dominant if temperature is higher. Many species are able to acclimate to the normal range of temperatures occurring below thermal discharges, so that when exposed to increasing water temperatures, their upper lethal limit is raised.

**D. Radioactivity**

Radioactivity means the phenomenon in which the nucleus spontaneously changes its nature slightly and emits a ray or particle of ionising radiation. Atoms of the same chemical elements may vary in the number of neutrons they have and are known as isotopes. Some isotopes are unstable and they seek stability by giving off particles or electromagnetic rays.

The environment is naturally radioactive, with radiation coming from outer space, from the earth, from the atoms within the body and from normal activities as burning fuel and cultivating soil. Sources of radiation from the ground are very variable regionally, depending on the geology. In general there is clear that some 88% of radiation received is natural origin. We have no control over it.

The major controllable sources of radiation to the general environment are those from nuclear weapons testing and from the nuclear industry.

The main concern over radionuclides is obviously the contamination of food chain, lead to the man. A small numbers of radionuclides tend to be of special importance. At high concentrations of radionuclides cause the acute toxicity.

The normal wastes released from nuclear power plants have not been show to cause harm to the environment or to humans.

**E. Eutrophication**

Eutrophication is the process which occurs in case a waterway become too rich in plant nutrient, especially phosphorus and nitrogen. The income of nutrients may be considered as artificial or cultural eutrophication if the increase is due to human activities, or natural if the rate of increase is caused by non human process. The majority of pollution nutrients enter watercourses and lakes in effluent from waste water treatment plants, in untreated sewage or from farming activities within the catchment. Industrial sources may be locally important, depending on type of industry. Food processing generally is likely to produce effluents containing high concentration of nitrogen and phosphorus. These are the discrete sources. Diffuse sources [rural sources] include from agricultural, from forest management and rural dwellings. Nutrients are lost in three ways:

1. by drainage water percolating through the soil leaching soluble plant nutrients,
2. by inefficient return to the land of the excreta of the stock,
3. by the erosion of surface soils or by the movement of fine soil particles into subsoil drainage system.
Eutrophication causes marked changes in the biota. The problems, which are caused by eutrophication is possible to summarize in three main areas:

- Problems associated with water treatment technology,
- Problems associated with aesthetic and recreational activities
- Problems associated with the management of watercourse and lakes.

Water monitoring

It is recommended to recognize nine distinct objectives of water quality sampling exercises:

1. to identify abnormal concentration,
2. to define the peaks e.g. dissolved oxygen,
3. to estimate mean concentrations,
4. to detect a trend or change,
5. to monitor a well-understood process,
6. to determine fair [and unfair] consent values for effluent,
7. to estimate a percentile,
8. to monitor for compliance,
9. to build up the picture of these process.

All these objectives can be summarized in three reasons for monitoring water quality:

A. to assess the state of rivers and the variability of quality,
B. to determine action if necessary to maintain and enhance quality,
C. to assess the effectiveness of such action.

2.3 Soil pollution

Monitoring of soil pollution is mostly focused on places exposed by industrial and agricultural activities. Whereas some parameters of air and water pollution are obtained from automatic and manually-operated monitoring networks, level of soil pollution is estimated from individual samples. To preserve comparability, the samples are collected from the same places. Examination of long-term changes requires expensive laboratory analyses. Regular measurements are usually complemented by data obtained from local project focused on dispose of soil contamination by disasters and waste dumps. The principles of evaluation are based on rules carried out by local and state authorities. Mostly, they determine maximum permissible limits of hazardous elements (heavy metals, inorganic and organic substances, polycyclic aromatic hydrocarbons).

Although primary measured data parameters do not require building of large database structures, they represent key factors in complex model analysis and risk assessment. Spatial data structures supported by the GIS together with measured data provide the base for geostatistical methods and dynamic modelling. The geostatistical methods are focused on spatial interpolation from point measurements. The analyses can be extended by using data from digital elevation models (DEM) and using functions of map algebra.
Dynamic modelling of soil pollution is mostly focused on accumulation, transport and interaction of contaminants in soils. Due to the heterogeneous environment, the model description is represented by partial differential equations. The lack of data to estimate all the parameters and initial conditions is mostly eliminated by simplified model description. The GIS is used like in modelling of air pollution for preprocessing and postprocessing. The solution of simplified models in the frame of GIS can be carried out in the programming environment of some GIS (ArcView-AVENUE, ArcGIS-ArcObjects). This way offers direct support of many function implemented in the GIS and remote sensing modules (Matějíček, 1996b, 2001).

2.4 Landscape protection

Urban landscape represents from the view of protection important element, which constitute living environment for urban population. Due to violation of former ecological systems and heavy antropogenic impacts, extra funds and projects are necessary to preserve existing nature. Conservation of flora and fauna include mapping of existing condition and monitoring short-term and long-term trends.

Information systems support evidence of land use and land cover. Except cadastral map, information about protected areas and natural parks involve other data sources for decision-making management. GIS together with remote sensing can map actual situation. Aerial photographs and satellite images help to identify and classify changes in the landscape morphology. Then all the data can be processed by the GIS. From the modelling perspective, landform changes can be explored by wide range of GIS tools (Matějíček, 2001b). In comparison with modelling of pollution, the methodology and modelling tools are focused on observation of complex natural, cultural and historical characteristics of landscape.

2.5 Waste management

The treatment and disposal of waste has developed from its early beginnings of mere dumping to a sophisticated range of options including today operations like re/use, recycling, incineration with energy recovery advanced landfill design and alternative technologies like pyrolysis, gasification, composting and anaerobic digestion.

Our aim, integrated waste management has been defined as the integration of waste stream, collection and treatment methods, environmental benefit, economic optimization and social acceptability into a practical system.

Waste is defined as thing, which the holder discards or intends to discard. Wastes are divided in many categories e.g.:
- Household wastes – means waste from private domestic accommodation, residential homes, universities and schools, hospitals and so on.
- Industrial waste – means waste from factories, premises use for supply water, gas and electricity or sewerage services.
- Clinical wastes – are wastes which consists wholly or partly of human or animal tissue, blood, drugs or other pharmaceutical products, needles or other sharp
instruments. Also any other waste arising from medical, nursing, dental, veterinary, pharmaceutical or similar practice.

- Municipal solid waste – there is common used for waste collected and disposed of by or on behalf of a local authority.

Except these wastes there is second group of wastes called hazardous waste. Hazardous wastes are defined like a waste that are or may be so dangerous or difficult to treat, keep or dispose of that special provision is required for dealing with them. They may contain substances which are corrosive, toxic, reactive, infectious, carcinogenic or harmful to human health, and they also may be toxic to environment. The OESD estimated of arising of hazardous wastes in European countries.

In the Czech Republic there is a project which deals with collection, management and recovery of municipal solid wastes. The project will determine the quantity, and composition of wastes. This data is of vital importance to the system as it enables the determination of the most appropriate waste management option for the types of waste managed.

In order to find the best solution for the waste management the project was subdivided in four basic parts:

- municipal waste quantity and quality,
- wastes collection system,
- recycling process,
- waste management mathematical model

A. Quality and quantity
The data concern quantity and quality of municipal wastes are basic data for next decision-making in waste management strategy. The strategy that has been developed is known as the multi-option strategy and it encompasses different ways of treating waste according to its recovery potential.

B. Collection systems
The recovery and recycling requires the effective collection system, one which reduces the level of contamination that would otherwise either prevent recycling, or generate a low quality end-product. Project provides and evaluates some option of collecting like e.g.:

- **Door-to-door selective collection**
  - Selective containers have been supplied to individual household for source segregation of recyclable materials, including paper, plastic bottles, and metal cans.

- **Eco-centres**
  - A network of eco-centres allows the effective coverage of a whole area. These sites are permanently manned and are used by local residents to deposit waste owing to its site or characteristic, cannot be removed by normal collection screw, such a garden wastes, wood, scrap metals, batteries, oils and light – bulbs.
  - Dustbin collection
- **Bring collection**
  - The citizen brings the re-useable waste to the place for collection and separation.

- **Mobile collection**
  Collection of separated parts of waste (garden waste, emitters, pharmacy products, bio-waste)

- **Stationary collection**
  - Collection keeps by dustbins, containers and special tanks.

C. **Recycling processes**

The aim of the recycling processes is a material recovery facility. The source of segregated wastes is processed for transportation to some recycling enterprises. The part of this sub-project is LCA. This is a tool to assess the use of resources and the potential environmental impacts associated with a product or process.

2.6 **Noise**

Whereas contamination of basic part of living environment (air, water, soil, landscape, waste) is perceived more intensively than excessive noise, the long-term effects of increasing traffic and large number of people living on small area are also very intensive factors affecting human health. A number of ailments (stress, neuroses, undesirable blood pressure changes) are manifested with a delay of several years. Information about noise levels, which are expressed by a number of parameters (equivalent noise level $A_{eq}$ for various built-up areas, daytime and nighttime; maximum and minimum noise levels) are obtained by measurements and models. The final estimates are put to the noise maps that are used for decision-making processes of endorsement and sanctioning.

A wide range of information about noise sources, urban environment and measurements has to be organized to carry out maps and models. However GIS offer a wide range of functions and spatial data management, many of data and models are managed in standalone systems. New GIS and their development tools offer great advantages in data management and integration with thematic urban maps and other data. For the present, new technology of data exchange and GIS development tools offer improvement of data integration and calculation methods.

2.7 **Other sources of pollution (biomonitoring, allergy, ...)**

There are many of other components in the frame of urban living environment that affect inhabitants. Some of them are included in various biomonitoring studies (occurrence of plant that affect allergy, monitoring of occurrences of Lyme boreliosis germs in ticks, occurrences of small rodents, accumulation of heavy metals and organic compounds in food chains). A number of data sources represent individual projects and grants supported by state and industrial authorities. The GIS by itself plays an integration link to a huge number of data formats and calculation tools.
3. Identification of model structure and parameters

3.1 Deterministic and stochastic identification methods

Identification methods include building of models by using natural laws or empirical experience and estimating of their parameters. Models, which can take advantage of natural laws, describe processes more comprehensibly. Considering to a complexity of environmental pollution, the base of models is described with the using of physical laws (transport and accumulation of pollutants). Empirical rules and observation are used to elaborate a number of other phenomena (heterogeneity of urban environment, anthropogenic interaction). There are no strict rules for building of models in dynamic modelling, which means that a few different models can be developed for modelling complex processes of environmental pollution. The next phase of identification includes setting of parameters in mathematical equations. Both the structure of a model and setting of its parameters need long-term observation of real processes. During the identification the behaviour of the model and measured data are repeatedly compared. The additional corrections are made to decrease the deviations between the model outputs and data of the real system. The procedure is illustrated in Fig. 1. The outputs of model $y_m$ and measured data $y_p$ are compared. The differences $e$ between the $y_m$ and $y_p$ results in the change of setting of parameters or furthermore the correction of model structure. This is arranged on the basis of model criterions. In case of the process, the inputs $u$ represent phenomena, which can be managed by decision-making process (regulation of urban emissions or transport). In the model, the input $u$ includes variables independent on model behaviour. The input variables $v$ mark superposition of random phenomena in the frame of the real process. If the random inputs are insignificant, the deterministic methods of identification are used. On the opposite side, the stochastic models and identification methods are needed (Ball, 1992).

3.2 Implementation of identification methods in the GIS

The GIS contain a number of identification functions, which are mostly focused on estimates of geostatistical models. Spatial interpolations use deterministic and statistical estimates of parameters. Regression and correlation models can operate in...
addition with spatial data. In spite of these tools, dynamic modelling of environmental pollution contains more comprehensive extent of functionality. The identification methods can be built in the GIS in case of smaller models. The larger the model the more specific methods are needed. In practice, there are not any general technique of identification of model structure and parameters in the area of environmental modelling of urban areas. Mostly, the identification is based on user’s experience and the GIS power to integrate wide range of data and corroborative functionality.

4. Simulation of models

4.1 Description of model calculation

The models represented by mathematical equations describe fundamental interactions of real processes. The examination of their functionality is carried out by numerical calculation of mathematical expressions on digital computers. A number of programming tools and programs have been developed during the last decades to simplify entry of mathematical expressions and their solution. For example, Advanced Continuous Simulation Language (ACSL) represents a system for numerical calculation of dynamic models. The dynamic of the model is calculated by a few basic elementary operations: integration and time delay, which are complemented with a number of algebraic relations. Especially, integration is used to realize accumulation in dynamic models. Time delay can support modelling of transport phenomena (Matějíček, 1995). Besides the ACSL, which represents the standard tool for calculation of dynamic models, there is a number of other system or languages that can carry out similar operations. They are grouped into many classes along the modelling branches (air pollution, soil and groundwater contamination, noise assessment, waste management). The individual software systems use various data formats that cause obstacles for data exchange and complex analysis. The recent systems represent complex tools that can support for example groundwater modelling (GMS- Groundwater Modelling System supporting both finite-difference and finite-element models in 2D and 3D including MODFLOW, MODPATH, MT3DMS/RT3D, SEEP2D, SEAM3D, UTCHEM and FEMEWATER), watershed modelling (WMS- Watershed Modelling System- for all phases of watershed modelling inclusive HEC-1, TR-20 and TR-55). Other examples could be illustrated by modelling systems for air pollution, noise assessment, landscape protection, ecological analysis or ecological modelling (Matějíček, 2002).

4.2 Implementation of simulation methods in the GIS

In addition to standard simulation systems, which support data exchange on the level of files or database connection, the GIS operate over a number of formats that support a region of space and time within which the variables can be simulated. Naturally, the GIS by itself is focused mostly on spatial analysis, which cause incompatibility with
most of standard simulation software. Often, the data have to be converted by additional tools.

The dynamic models are divided into a number of groups (continuous or discrete in space or time, linear or non-linear, deterministic or stochastic). Each group requires different data structures and algorithms for its calculation. The GIS supports partially simulation of some dynamic models. For example, the GIS Arc/INFO can solve diffusion, which is implemented in the Grid module. The GIS Idrisi is extended with a number of modules, which support dynamic control structures that allow the development of temporal models (such as cellular automata, growth and change analysis). Besides standard functions, almost each GIS contains programming tools that enable implementation of own numerical algorithms for calculation of environmental models. Because complex models have considerable demands on efficiency of numeric calculations, models are used to be solved outside of the GIS environment.

4.3 Display and visualization of model outputs

Most of the simulation systems is focused on numerical calculations. Visualization is supported on the basic level. Each system contains specific functions and different user environment. The GIS offer standard environment and the possibility to extend graphic outputs with spatial data, images and photos. In addition, all the graphic output can be accessible through the Internet by the GIS browsers (ArcExplorer, CAD tools).

5. Verification and validation of the model outputs

5.1 Interpretation of model outputs

Any representation of a real system requires the development of an abstract model that is built by a set of basic constructs or variables and language for manipulating them. In environmental modelling developed by dynamic models, the abstract models are based on mathematical expressions that represent interactions of environmental processes. During the validation and verification, the structure and measured characteristics of the real system are repeatedly compared with the model structure and simulation outputs. In GIS, the abstract model rests mainly on the representation of geographic features by data connected to layers of points, lines, polygons, grid cells, images and photos. The processing of subsequent data with functions represents next steps, which is mostly focused on spatial interaction and transformation of original objects. In spite of that verification and validation are terms used in the context of dynamic modelling, the similar evaluations in the GIS environment means how accurate and complete spatial data are (location, shape, generalization and topology).

Verification and validation from the perspective of environmental dynamic models built inside the GIS associate both the time and spatial aspects. Especially in the urban areas, the complex spatio-temporal analyses represent multilateral problems.
Considering to the heterogeneity of modelling systems, verification and validation are often pursued separately for individual components of environmental pollution. Unfortunately, the multidiscipline orientation of the complex environmental models overtaxes the area of individual research specializations (air, water and soil pollution; landscape protection; noise assessment; GIS; dynamic modelling), which opens new research subjects (Costanza, 1991).

5.2 Error propagation and data accuracy

The difference between the spatial objects and time dependent phenomena of real systems has been caused distinct approaches in error assessment. The dynamic models generalize real systems through the fundamental and insignificant phenomena relative to the studied subjects. In accordance with this, discrete or continuous, and linear or non-linear models are used. The spatial analysis generalizes the real object in view of the scale. Similarly, this indicates the use of the vector or raster spatial data format, pertinently triangulated irregular networks-TIN for digital elevation models or other spatial formats. Despite of various levels of data accuracy, the complex analysis needs compatible data formats to analyze interactions of environmental pollution together with urban areas. One of ways represents cell based modelling, which can assess each cell in the grid by weighted factors of individual environmental processes. The schema in Fig. 2 illustrates data flows, which start from data collection, model development and basic evaluation to model simulation and statistical analysis. The final analyses are carried out in the GIS. In addition to a number of spatial analyses, the cell based modelling can be used for complex analysis of individual data outputs. In spite of the transparent data management in the form of the reclassified regular grids, the

![Fig. 2 Cell based modelling of environmental pollution in the urban areas.](image-url)
dynamic of phenomena of environmental pollution can be studied by restricted way. The cell based modelling, which is mostly used for simpler data assessment, cannot provide more complex time dependent analyses that are carried out by simulation of individual dynamic models.

6. GIS data structures

6.1 Spatio-temporal description of environmental pollution

Considering to the structure of data and functions that deal with the environmental pollution, the research requires spatio-temporal data management (Matějíček, 2002). New technology in the area of computer information systems is facilitating the advance in the GIS, which is primary focused on spatial data management. But not only technology can contribute spreading of the GIS. The economic factors and wide versatility, especially personal computers, move the GIS to the leading role in the area of information systems, which can support spatial analysis and partially dynamic modelling of pollution in urban areas.

The description of data structures suitable for the support of analysis and modelling pollution in urban areas from the GIS perspective is still influenced orientation of the GIS to assist digital mapping and spatial analysis. In spite of new software technology (ArcGIS objects, interconnections to the relational database systems, multimedia extensions), most of the GIS include basic support of dynamic modelling (calculation of diffusion in the GRID module of the ArcINFO GIS, cellular automata in the Idrisi GIS). On the opposite side, a wide range of data formats used by modelling software causes great obstacles for complex analyses of results. From this point of view, the GIS with their database interconnection represent an important step of integration and standardisation of data formats.

Among the basic spatial data formats used in the GIS to display urban objects and to analyse pollution phenomena belong raster and vector data formats. The principal difference illustrates Fig. 3 (an example of concentration of pollution in an urban area). The first schema is represented by the vector format. The contour lines built by polygons mark individual levels of concentration. The polygons are formed by closed chains of lines. The accuracy of their coordinates depends on the number of digits. Each place on the map belongs to the certain level. The remaining schemas represent various levels of air pollution by the raster format. The
Modelling of environmental pollution in urban areas with GIS

Accuracy of concentration in each place depends on the resolution of the raster grid. Other factors represent data demands. Whereas the vector formats are stored as positions of line features extended by information about interconnection and topology, the concentration represented by the raster formats must be stored for each cell. But in case of the regular grids, there is no need for information about interconnection or topology. The more the cells the higher capacity of a storage media is needed.

The spatial data formats in the GIS represent objects or phenomena in the reality (buildings, streets, administrative units, air circulation, concentration of pollution). The primary aspect covers the affiliation of the real object and phenomena to the GIS features represented by the vector or raster formats. Generally, the real objects can be divided into discrete and continuous. The discrete objects have definable boundaries (buildings, roads, rivers). Continuous objects (field, non-discrete or surface data) represent mostly phenomena that progressively vary as they move across a surface (elevation, diffusion and advection of pollutants from sources measured by their concentration). For representation of many real objects and phenomena, the boundaries are not clearly continuous or discrete. A continuum is created in representing of object features with the extremes being pure discrete and pure continuous (soil types, boundaries of ecosystems). Due to determined boundaries and irregular shapes, the discrete objects tend to be represented by vector formats. Whereas the continuous objects or phenomena, which are going to be analyzed and integrated together with other data, are transferred to raster formats.

Among a number of other GIS formats, the triangular irregular networks (TIN) are used for representing of surfaces as contiguous non-overlapping triangular faces. To compare to the raster representation, a variable point density in areas where the terrain changes sharply yielding an efficient and accurate surface model.

Modelling of environmental pollution is usually processed on regular grids. So, the raster formats are used for spatial data management of dynamic phenomena. The outputs of simulation can be consecutively converted to the vector formats or the triangulated irregular networks to display distribution of pollutants in the most suitable form. In spite of that, many of simulations operate with the regular grids. On the opposite side, data from monitoring systems, which are collected by stations, form an irregular point network. The intermediate points have to be estimated by interpolation methods that must reflect heterogeneity of urban areas. These data are used for setting of initial conditions and parameters of dynamic models concerned on mass accumulation and transport. Thus, the environmental modelling and representations of real objects and phenomena must operate with a number of dissimilar data formats. But the GIS in the frame of data management can execute conversions and realize other transformation (map projection, data exchange with other information systems).

6.2 Data tables and relationships

The graphic elements in the GIS in addition to the position and shape contain other information in the form of attributes. The attributes represent other characteristics
of graphic elements or can provide relationship for other data. Fig. 4 illustrates a point theme together with its attribute table. Each graphic element shares one row, which contains other data (time series of concentrations, specification of a source of pollution). In addition to these data, attribute tables can contain columns that enable joining other data from external tables. The connection is managed through the columns that contain identical items in each row of the attribute and external table. Most of the GIS transmit relationships among data tables to the relational database management systems (RDBMS). These systems (Oracle, SQL Sever) have been representing standard tools for data management for a few decades. Their capabilities also enable better data exchange and connectivity to the large information networks.

Spatial information, which is mostly arranged in the frame of vector themes (point, line and polygon elements) or raster and TIN layers, contains information about the position and shape of sources of pollution, levels of contamination, etc. Data, which specify for example time series of concentrations and amounts of pollutants, are joined. They are mostly obtained from monitoring systems or model simulations that are used to store measurements and model outputs in different data formats (database or binary data, ASCII, XML and other formats). So, databases and standard data formats like XML can represent flexible environment for data management. The GIS participate in this structure together with dynamic models as a case system for spatial analyses and model simulations.

6.3 GIS’s analysis

The GIS capabilities originally focused on the digital map composition have been extended during a few last years. Now the GIS offers a wide range of spatial analyses, which include spatial data conversions, query tools, cell-based modelling, spatial interpolations, network analysis, visualization in the space, map compositions and display, multimedia extensions, etc. In addition, the most of the GIS contains programming environment that can extend and complement existing tools. Due to the heterogeneity of urban areas, the GIS together with its spatial analysis and data management represent versatile tools that can support a wide range of analyses of urban environment.

The dynamic modelling of environmental pollution overtaxes calculation possibilities of the present GIS. In spite of new technology of sharing data and programming code in the frame of the component object model (COM, for example implemented in the ArcGIS as the ArcObjects), the numerical algorithms for calculation of trans-
port of pollutants must be implemented individually. Also, dynamic modelling has been growing up separately for a long time. The most of simulation systems operates with specific data structures and different functions for sharing of spatial information. Moreover, each group of environmental models has built individual tools subordinated to the purpose of its field of study. All these factors represent obstacles for building of common data structures and functionality together with the GIS. Generally, a few ways can accomplish integration of dynamic modelling and the GIS. The principal schema of interconnection of dynamic modelling and the GIS illustrates Fig. 5.

The low level of integration represents data exchange between the GIS and the standard simulation system, which substitutes a prevalently used standalone programs. The data must be converted by individually developed procedures. This way can take an advantage of the existing software, but the data exchange degrades efficiency of complex analyses.

The obstacles with the data exchange can be overcome by rewriting dynamic models describing the environmental pollution in urban areas into modules, which can directly use data from the GIS data structure. There is no need to carry out new mathematical models in the frame of the development. The basic laws governing the motion of pollution in air, soil, surface water and groundwater have been elucidated in the intervening years. The key is to implement the principal phenomena in individual cases of urban areas with the use of software development tools (ArcObjects in the frame of ArcGIS, AVENUE in ArcView). The description of dynamic phenomena can be made by lumped or distributed models as described by Maidment in the book edited by Goodchild (1996).

The lumped models are used to be described by ordinary differential equations or in a simpler case by algebraic relations. They mostly describe changes of separated processes, which are spatially allocated into points, lines or polygons. These elements belong to the vector data structure in the GIS. The changes of state variables are stored in attribute tables or joined external tables. The lumped models that for example describe mass or energy accumulation in separated spatial elements can be described by ordinary differential equation:

\[
\frac{dS}{dt} = \sum (f_{in} - f_{out})
\]

where \( S \) is the amount of the mass or energy, \( f_{in} \) and \( f_{out} \) represent input and output flows. The flows are time dependent functions. In case of lumped models, the structure, parameters and initial conditions of each equation must be set individually for each element. The mathematical model (1) has to be transformed to the numeri-
cal model that provides an approximate solution. The first order approximation for lumped models can be expressed by the formula:

\[ S_{k+1} = S_k + \Delta t \sum (f_{in} - f_{out}) \]  

(2)

where the \( S_k \) represents time sequence of values calculated in the time interval \( \Delta t \). The initial value \( S_0 \) must be estimated by field measurements. In practice, the more complex formulas are used for more precise calculation. As an example, modelling of pollution of water sediments in a few layers is illustrated in Fig. 6 (Matějíček, 1996b). The arrows mark measurements of concentration of the pollutant. The areas are represented by polygon elements in each layer. The vector formats are used for data management inside the GIS. \( x \) marks amount of the pollutant, which is transferred to the estimate of concentration. The attached schema illustrates the process of solution of the dynamic model in the environment of the programming tools AVENUE in the GIS ArcView.

By contrast to the lumped models, the distributed models operate over a continuous space in one, two or three dimensions. Generally, they are assembled with a few basic phenomena: diffusion, advection, chemical interaction and description of the source inputs of pollutants. The fundamental two-dimensional system can be described by a partial differential equation:

![Fig. 6 Simulation of pollution of water sediments developed by programming tools in the GIS.](image)
\[
\frac{\partial c}{\partial t} = \frac{\partial}{\partial x} \left( D_x \frac{\partial c}{\partial x} \right) + \frac{\partial}{\partial y} \left( D_y \frac{\partial c}{\partial y} \right) + K_x \frac{\partial c}{\partial x} + K_y \frac{\partial c}{\partial y} + f^{ch}(t,x,y) + f^i(t,x,y)
\] (3)

where \( c \) is concentration of the pollutant. \( D_x, D_y \) and \( K_x, K_y \) represent diffusion and advection in two directions (x, y). \( f^{ch}(t,x,y) \) is the term that contains all the physical and chemical interactions. The term \( f^i \) describes the inputs or outputs of the pollutant. The terms \( f^{ch}, f^i \) and all the parameters can be time and spatially dependent. The example of the two-dimensional model of transport pollution can be extended to three dimensions or simplified to one dimension. The more the dimensions, the more complex the system, which causes larger amounts of calculation. Like in the case of the lumped models, the formula (3) has to be transformed to a numerical formula. An example of the explicit numerical approximation is expressed by the formula:

\[
c_{i,j,k+1} = \tau \left[ D \frac{c_{i-1,j,k} + c_{i,j-1,k} - 4c_{i,j,k} + c_{i+1,j,k} + c_{i,j+1,k}}{h^2} + K_x \frac{c_{i+1,j,k} - c_{i,j,k}}{h} + K_y \frac{c_{i,j+1,k} - c_{i,j,k}}{h} + f^{ch}(t) + f^i(t) \right]
\] (4)

where \( c_{i,j,k} \) are the concentrations in the cells of the regular grid, which are indexed by \( i \) in the direction x and by \( j \) in the direction y. The index \( k \) refers to the value of the discrete time, which is divided to the time intervals by the time step \( \tau \). The time sequence is similar to the sequence of the state values of the lumped models. The constant \( h \) is the distance of the cell centres in the regular grid. A number of other more accurate approximations of the model (3) can be carried out. The more precise the approximation, the more calculations is needed, which can cause in the case of two or three-dimensional models long simulation times. Considering to the regular grid, which is formed by the formula (4), the raster format is preferred for data management in the frame of the GIS. As an example, the general schema in Fig. 7 offers the basic overview for a development of the three-dimensional distributed models. In case of the modelling of air, water or soil pollution, the three dimensions are used to be reduced to two-dimensional layers, which can be manage by the GIS in a more effective way.

![Fig. 7 General schema for model description of the air pollution.](image)
Generally, in case of the simulation of transport phenomena, the distributed models require more calculations than comparable solutions by lumped models. Of course, the description of both the models types proceeds from different assumptions. The both described concept of lumped and distributed models can be implemented in the GIS. The efficiency of the simulation depends on a conception of programming tools. The most of the GIS contains various types of the macro or script languages, which are primary designed to automate sequence of procedures mostly offered by a user interface. This can save time in case of repeatedly processed operations. But in case of calculations of spatio-temporal models, the more efficient environment with basic programming structures is required. Fig. 8 demonstrates integration of environmental modelling in urban areas in the frame of the GIS ArcView (Matějíček, 2002a). The simulation system is developed by the AVENUE programming language. In addition to the extension for spatial, 3D and statistical analyses (Spatial Analyst, 3D Analyst and scripts for statistical analysis), the simulation of dynamic models can process lumped and distributed models. Due to the structure of the AVENUE and its efficiency, the system is focused mostly on the lumped models. Simulation tools operate directly with the attribute data, which can be used for other data management and visualization of model outputs in the GIS, Fig. 9.

The integration of simulation together with the GIS tools gives possibility to carry out the complex analysis in the urban areas just in the GIS. So the wide range of simulation tools for a relatively different field of study can be managed more efficiently. This complex assessment of different phenomena, which affect living environment in urban areas, is able to support decision-making processes in a more complex way. On the opposite side, the present GIS by itself contains just a small part of modelling tools that are dedicated for the direct support of complex environmental problems. Many of analyses and functions must be still added to the GIS to approach the more complex analyses.

Fig. 8 Environmental modelling and spatial analysis of pollution in urban areas.
Besides the dynamic modelling in the independent simulation systems or integration of the dynamic models into the GIS, the GIS programming libraries can be used to developed standalone software applications. This is one of the most efficient ways, which offers the direct solution of dynamic models inside programming modules supported by the GIS functions.

6.4 Object oriented data structures

In a few last years, the new trends can be observed to put together data and methods that supply their management. The programming structures with their data and statements, which are encapsulated, inherited and polymorphic, form objects. A number of definitions and specific terminologies exist in the frame of object oriented programming languages and development tools (Spanou, 2000). From the point of view of the GIS, it means that spatial data should be given together with the appropriate functions or methods. In practice, there are various levels of the support of object programming styles in the present GIS. The sophisticated structure is developed in the frame of ArcGIS, which is called ArcObjects. Also, this building set of many objects give opportunity to develop a strong application for modelling of environmental pollution. Despite of a low level support of dynamic modelling, the collections of objects enable flexible construction of a wide range of tools that can assist to solve complex environmental models.
7. Modelling of environmental pollution: extensions

7.1 GIS and Remote Sensing

In additional to the presented analysis of environmental pollution in urban areas, the GIS together with remote sensing offers far more tools for decision making processes in the frame of urban areas. Satellite images or aerial photographs offer information about an actual condition of the urban surface. Moreover, they can provide other supplementary information that is used for the identification and location of environmental accidents. The remote sensing, which is primarily focused on the acquisition of data about objects or scenes by sensors that are far from the earth surface, also includes processing and analysis of data from aerial photography, satellite imagery and radar. The electromagnetic energy is transmitted through space. The remote sensors record specific wavelengths of the electromagnetic spectrum. The various types of a land cover (rock types, water bodies, forests, agricultural areas) absorb different portions of the electromagnetic spectrum, which is used for analyses of remotely sensed images. The classification methods and comparisons of images in time sequence can give valuable information about the consequence of the environmental pollution.

Fig. 10 Satellite images from the LANDSAT 7 with bands 1, 2 and 3 (Prague, May 2000).
Fig. 10 and Fig. 11 illustrate the satellite images from the LANDSAT 7, which display the area of Prague. In addition to the composition of the visible portions of the spectrum (bands 1, 2 and 3), the composition with the reflective-infrared spectrum is demonstrated (bands 1, 2 and 4). This composition offers better discrimination of the water content in soil and vegetation. This emphasizes the condition of vegetation, which can indicate in the time series of images affects of urban environmental pollution.

7.2 Optimization analysis

The methods of the optimization analysis in the frame of dynamic environmental models include modifications of model parameters to improve the behaviour of the models (minimization of pollution by building barriers, different locations and regime of sources of pollution). The special discipline forms methods on the base of linear programming, theory of graphs and stochastic processes described by Markov chains.

As an example, the optimization of transport of municipal waste in the urban area is presented. The information systems, especially GIS, can optimize waste streams on
the base of economic calculations complemented by environmental benefit and societal acceptability (Matějíček, 2001c). The GIS can support optimization from sharing data among various programming systems to fully integrated applications. Apart from finding efficient travel routes (low costs, rate of disturbance and environmental risk), which can be carried out by the GIS, the optimization models based on linear programming enable more sophisticated descriptions of decision-making processes. For example, the transport model, which minimizes the total cost, is described by the criteria function:

$$f(x) = \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}$$ \hspace{1cm} (5)

subject to:

$$\sum x_{ij} = b_j$$

$$\sum x_{ij} = a_i$$ \hspace{1cm} (6)

for all \(i\) and \(j\), where \(c_{ij}\) is the costs per a unit of the waste material, which is transported from the \(i^{th}\) source to the \(j^{th}\) processor, \(x_{ij}\) represents the amount of the waste material, \(b_j\) is the capacity of the \(j^{th}\) target side, \(a_i\) is the capacity of the \(i^{th}\) source site. An implementation of the transport model described by equations (5) and (6) can be developed with programming tools of the GIS. In addition to the model implementation and the use of standard GIS functions, the model application can be shared through the computer network. Fig. 12 illustrates a case study of the GIS project in the environment of the Internet.
8. Conclusion

The synthesis of dynamic modelling, the GIS and optimization methods in the frame of environmental modelling of pollution in the urban areas has the potential to create more sophisticated and complex analyses that can support research and improve the decision-making processes in practice. To accomplish all the requirements of complex environmental analyses particularly focused on environmental pollution, a number of dynamic models and optimization methods have to be translated into the GIS. On the opposite side, the GIS is still oriented on spatial data management need more extensions and more efficient programming tools that can support spatio-temporal analyses and simulation of larger models. A few ways has been presented in the previous parts to accomplish an overcome of these limitations.

The lowest level represents sharing of data between the individual modelling systems and the GIS. This way has been improved in the past years by technology of data exchange and linking. For the present, sharing of data can be used for all cases of the connection of individual expert systems together with the GIS. But, it includes series of obstacles that consist in the data format incompatibility and the needs of additional programming. Moreover, some expert systems for modelling of pollution phenomena do not allow an export of outputs in the formats that can be included to the GIS’s layers.

The way of the development of the own simulation tools for calculations of the lumped or continuous models requires expert knowledge of programming in the frame of the GIS development tools (ArcGIS’s ArcObjects, ArcView’s AVENUE). On the opposite side, environmental analyses can be carried out in the frame of one system (GIS), which offers more complex analysis. Also, there are no additional needs in the extended GIS to standardize data incoming from other systems.

The both previous ways of integration of the individual simulation systems and the GIS require the presence of a GIS computer program. The GIS by itself contains much more capabilities that are not necessary for environmental modelling. The most efficient way from this point of view represents the standalone application built with the using of the modelling and GIS software tools. But the needs of the programming experience and expert knowledge in the field of the multidisciplinary research require wider cooperation.

The main attention in this contribution is dedicated to the present situation of environmental modelling in the frame of the GIS dedicated to the urban areas. In spite of the research progress, the GIS are still concerned more on the digital mapping applied to a wide range of disciplines than on the calculation of environmental models, the present development systems (for example ArcGIS’s ArcObjects) offer an implementation of the modelling of diffusion and advection phenomena in the continuous models, accumulation and transport in the lumped models and the optimization methods. Considering to the spatial data management, which can be extended by connection to database systems, the GIS can support complex analyses and risk assessment in a more sophisticated way. In practice, the dynamic environmental models integrated to the GIS together with maps, aerial photographs or satellite images offer
more intuitive ways for interpretation and presentation of environmental pollution in the heterogeneous urban areas. This can also contribute to the decision-making processes on the level of state authorities and local agencies.

Acknowledgements

The presented models were developed in the frame of the project supported by Royal Society in London and grants focused on environmental studies sponsored by Charles University in Prague and the Ministry of protection of the Czech Republic. The final outputs were carried out in the GIS Laboratory at the Faculty of Science, Charles University in Prague. Experimental data and digital maps used for simulation of environmental models are partly administrated by the Environmental Department of the Prague City Hall.

References


