PROBLEMS AND SOLUTIONS OF FIELD SCALE AGRO-ECOLOGICAL DATA ACQUISITION AND DATA INTERPRETATIONS IN AGROINFORMATICAL DOMAIN

J. TAMÁS
e-mail: tamas@gisserver1.date.hu

Department of Water- and Environmental Management, University of Debrecen, H-4032 Debrecen, Böszörményi út 138, Hungary
(phone / fax: +36-52-508-456)

(Received 3rd Nov 2002; accepted 26th June 2003)

Role of agro-ecological data sources for sustainable development

Agricultural production is an activity of high economic and environmental risk, nowadays especially. Handling of widening scientific knowledge and up-to-date technological appliances needs rethinking of foregoing theory and practice; according to some researchers, total paradigm change is needed. The “sustainable development” theory of Agenda 21 made up by the Rio Summit (1992) means long-term tasks for all branches of economy, so as for agriculture. As a result, the change of land use planned from economic and environmental points of view would realign the overall view of the Carpathian Basin for centuries in a territory where the ratio of agricultural regions is one of the highest (72%) in Europe. Risks of decisions refer to any period or area can be decreased by having more and more data and analysed information. Agricultural activity, as an open ecological system, has a hardly elaborative data requirement even by the present information technological devices [12] About half of the input data interfering the results are changed during a fragment of the phenological cycle, that can not be described exactly by such widely used ecological models as CERES [25] or CROPWAT [44] during particularly extreme changes validated for Central-European model border conditions. These results in more data need in space and time, which has properly been satisfied so far by the development of digital sensors, but the practice is not able to keep up with the manipulation integration and interpretation of the collected data. Similar process occurs on the field of data mining technology performed on Internet, where effective solution of information production still has to be waited for. Seemingly, this causes contradiction, the decision maker uses smaller and smaller part of the information available, but this data amount – in absolute sense – grows at a highly increased rate in time. During many decisions the incomplete interpretation assumes the reliability of data or their interval and value of uncertainty are not known. As a result, decision makers have no numerical risk values actualised according to the information built up from the data, thus the farmer dares to undertake lower and lower risk level in the smaller decision space.

How one can escape from the trap of this lack of data and data dumping that sometimes comes up at the same time in destructive cases? The key element of the decision-making process is improvement of the interpretation techniques that must be optimised during the whole data flow.

ISSN 1589 1623

Optimisation of data mass, data quality and set of information effectively support the sustainable development at the given decision level, where the decision maker knows the levels of uncertainty of the data, the error spreading processes and the undertaken decision risk in space and time.

In decisions, the economic market information was dominant up to the mid 1990's in the Hungarian industrialised agriculture. The structure of ecological-environmental information data was determined by this dominantly agrochemical point of view in Hungary. From amongst the ecological and economic environments in spite of the renewing economic crisis the European Common Agricultural Policy is calculably stable, in which the agro-ecological and rural development elements have been upgraded. The ecological environment keeps on changing capriciously, in fact the rapid urbanization and the conversion of the environment causes more often natural hazards (draught, floods, exceeds waters etc.) with significant damages. In the agro-decision space the economic and classical crop production and animal husbandry questions mean smaller, while the ecological tasks bigger role [36]. Determination of potential of natural resources, monitoring of the changes is an agro-information task even nowadays in the ideal several-meter spatial resolution as well as in the several-hour temporal resolution, whose mass distribution is a big challenge and high-tech task for establishment of an information society.

Therefore amongst the interpretation techniques the possibilities and limitations of the agro-ecological digital data acquisition modelling that is prevailing for sustainable development will be overseen.

Evaluation of Hungarian natural resources

In the document Agenda 21 a proposal was born to establish the conditions of Sustainable Agriculture and Rural Development (SARD) that aims the establishment of food production and food safety as well as protection of the resources at the same time. Previously, in Hungary after the world energy-crisis of the 1970's several research and governmental programmes were started to evaluate the volume and utilization of natural resources used by agriculture [9, 30]. Results of the project called "Agro-ecological Potential of Agriculture on Millennium" led by Láng et al. (1983) [28] were systematised into a professional system by the *Agro-Ecological Integrated Information System (AIIR)*. AIIR is a professional system that manages and builds on settlement, soil, meteorological, land use and crop production data.

Its main advantage was that resources were evaluated by up-to-date, model methods. During the that time of "industrial-like" production purposes, where the goal was the hospitalisation of productivity (y), the limited possibility of sustainability of resources was pointed out and the questions of yield safety – risk as well as production site conditions were discussed.

According to the model factors of the methodological researches led by Harnos [15] the stationary state of the soil characteristics (x) in time (t) were assumed, where t is only one phenological cycle in evaluation of sustainability, while x can change depending on time [$x(t)$]. A group of climatic factors can change rapidly and randomly (ξ), while the other group applied in agricultural technology (u) containing all elements of production (seeds, nutrients, plant protection, soil cultivation, irrigation and crop rotation etc.), can change in every phenological cycle. The effect of the production site to the yield (η) was described by the following equation:

$$\eta = \eta(x, \xi, u).$$

In the sustainability model it was interpreted that productivity of a given x production site is not decreased by the u-type agricultural technology at t+1 time. That is, quantity of yield under the actual agro-technological solution will not decrease in a further production cycle.

$$y_{\text{act}}(t) \geq y(t+1)$$

Productivity is expressed by $E(\eta(x(t), \xi, u))$, where E is the deviation caused by the weather, whose value can be characterised by the $F(y) = P(\eta(x, \xi))$ distribution function, where $F(y) = P(\eta(x, \xi)) \leq y$ means that yield of the reference plant (y) will not be higher than the average yield expressed in t/ha. In the relation the P random variable is $0 \leq P \leq 1$. The approach above gave a possibility for the probability establishment of yield loss (p_v) as well. Hindrance of introduction of the theoretical model was that the analytical form of the $\eta(x, y, u)$ function describing the connection amongst yield–production site–agricultural engineering and the $x(t+1) = g(x(t), u(t))$ function describing the state change of the production site were not exactly known.

In practice the studies were focused on:

- genetic and agricultural engineering development and
- stochastic change of weather,

where the county units were used as spatial references, under the actual technical conditions.

Agricultural technology is expressed by the change in time of η. Accordingly the change of average yields in time (y_1, η_2) can be expressed by the $\eta(t, \xi) = y_1(t) + \eta_2(t, \xi)$ potential function, where $y_1(t)$ describes the stochasticity of genetics and applied technology, while η_2 is that of the weather. During the almost two decades passed since these studies the role and function of agriculture and its methods of research have been changed. The need of data and the methodology of evaluation of the regional and local ecological information systems build on large-scale field information systems are upgraded. Unfortunately, its introduction to practice delays because of lack of intellectual and financial capitals supports.

Soils and land use

Evolution of soils is affected by geological, climatic, topographical and biological features as well as by the age of soils. Analysing soil evolution factors, effects of human activity on soil cannot be neglected. This effect has been particularly intensive during the last several hundred years. On the one hand, this human activity resulted in the improvement of productivity of soils, while on the other hand, in certain regions, it accelerated soil degradation processes. All of the above-mentioned soil factors exerted their effects together in the Carpathian Basin and their interactions determined the form of appearance, the physical, chemical and biological features of the certain soils [41]. Processes in the soil represent opposite impact pairs that are in dynamic balance in space and time [1]. These balance processes can shift to one or the other direction, strengthen, change periodically in time or may have shorter periodic effects that can be temporary or permanent.

On the AGROTOPO 1 : 100 000-scale digital soil map of Hungary, Várallyay [50, 51] distinguished 3310 polygons based on 10 soil parameters. In case of 1 : 10 000 – 1 : 25 000 field-scale levels the characteristic values of the heterogeneity of spatial

ISSN 1589 1623

patterns grow at a highly increased rate. With increase of the spatial resolution the standard deviation of attributive data increased as well. In this dimension the digital topographical data was partly processed, which is the main data source for survey of soil resources and for performing an agro-ecological model. At determination of soil characteristics on field, the representativity of parameters got for a general purpose soil survey are interpreted by Webster [52] as how much it can explain from all the variances of the data. He asserts that in case of soil-physical characteristics about half of them, while in case of some soil-chemical characteristics less than one-tenth of all variances are revealed by the survey. It must be overcome by the soil-characteristics estimation based on environmental correlation.

From amongst the soil characteristics the biological, chemical and physical parameters can be measured in growing uncertainty order, respectively. Among the soil-physical characteristics, the water- and heat management values of the same sample can be differentiated at several thousand-fold measured either in situ on the field or in the laboratory.

Uncertainties of soil-chemical measurements are grown in order of magnitudes by preparation and extraction procedures prior to measurement, in comparison to the interval of value of analytical device measurements. The up-to-date field sensors and remote sensing data significantly lessen the volume and deviation of sample preparation and of laboratory errors. While ten years ago several weeks were needed from soil sample taking to the evaluation of the result, nowadays this can be lessened only to some hours. The measurement limits for example in agro-chemistry in case of field devices decreased to ppm, while in case of measurements in a laboratory to ppb level value.

Digital technology defined numerically the experimental fact that in the Great Plain regions of Hungary the pedological changes are significant even in vertically in sub-meter level and due to the intensive land use they take place much faster than it was estimated previously. This fosters new research directions. Applying simulation models can help in determination of direction and volume of the processes. A group of wide-spread used models contains such regional, for example erosion models like USLE and WEPP, while their other group deals with the nutrient management of the given site, e.g. SOIL-SOIL-N or with water movement at point scale [24, 35]. Some of these models are for research, need a big volume of input data and provide a very detailed process description, while the others are practice-orientated, robust models.

Water sources

After the Netherlands, the second largest area being threatened by floods and exceed waters can be found in Hungary. Frequency of extreme values in the statistics of hydrology has grown trend-like during the past decades [34]. The extreme changes in the water source together with the permanent contamination burden increase the variability of the quality of waters as well [38, 39].

Runoff of exceed waters closed out from flood areas, the volume, period and frequency of inundation are consequences of series of accidental hydrological events and phenomena [40]. Probability of floods in space and time can be basically determined by two ways: on the one hand, as volume of frequency values of inundations, on the other hand, as size of factor maps of exceed waters. Besides,

ISSN 1589 1623

combination of these two can occur, which means the subjective correction of the first method by using the second one.

With re-classification of the digital maps of such characteristics that take part in formation of exceed waters and show relative spatial stability (for example aquifer hydraulic conductivity and maximum storage capacity of the aquitard, convexity of the micro relief, critical possible depth of the underground water, land use), thematic maps can be prepared. From their subsequent overlaying an exceed water risk map can be prepared [4].

There are several possibilities for spatial distribution and interpretation of the infiltration factor. The simplest way would be making an isometric map if the co-ordinates of the sample taking sites were known. But this solution might come up only if we would like to interpret a soil-genetically or soil-physically homogeneous area. Then by choosing the proper interpolation technique a continuous surface can be modelled, that is suitable for interpretation to "all points". But the heterogeneous regions have several anisotropy, that are anisotropies having such impacts that significantly change the regularities of spatial development of the infiltration factor. These are mainly the factors that play role in formation of the soils (for example the topography – i.e. the deep areas have low water conductivity).

Therefore, if a map that pictures a detailed physical kind of soil is available whose profile analysis data (for example the mechanical structure) are known – that are base of preparation of the map –, then through the revealed relations thematic maps displaying the hydraulic conductivity of the bigger regions can be prepared. These maps are true with the given input structure and boundary conditions. The spatial pattern of the change of the accumulation conditions overpasses the speed of change of soil characteristics in time [45]. Mathematical survey of time series analyses is indicated a fast developing field.

Climate

Material of knowledge of Hungary's climate potential is as revealed as that of soil sources. Climate potential for country size and meso-regions were determined to study the maximum yield [2, 42, 48], which provides spatial modelling possibilities. Based on their study, a part of the total radiation energy that reaches the plants (Q_0) is photosynthetically active energy (Q_p), whose certain part (ε) is used by the plants for biomass production (Y_b)

$$Y_b = \frac{Q_p}{Q_0} \cdot \varepsilon \, .$$

Burgos [5] concludes the value of the maximum possible utilisation of radiation from the estimated value of the energy reaches the Earth. Based on it, Varga-Haszonits et al. [49] calculated the value of ε to 22–23% for Hungarian conditions. Relying upon the Campbell method [7] performed also by him, he compares the energy of photons relate to the medium wave of the photosynthetically active radiation with the amount of energy bound chemically in 1 mol material. Its value is about the same: 22%.

Climate exerts its effect – besides the energy conditions – through the fluctuation of water resources. During modelling the Aridity Index that expresses the rate of potential evapotranspiration and precipitation during the vegetation season, has an average value of 1.6–2.8 in case of winter wheat in the 1951–1991 period of time.

ISSN 1589 1623

According to the time series analyses calculated by Harnos et al. [16], for the period of 1951–1983 there are 3–5 draughty years in every 15 years, that cause yield loss of higher than 5%. From amongst the three cases the yield loss once is 10–15%, once is over 15%.

Eight research stations were included in the above-mentioned survey. But the field scale forecasts are very uncertain. The present destiny of field measuring stations of meteorological service is not able for supervised teaching of satellite multispectral images. They cannot supply continuous real-time data for field-scale water- and heat management analyses but at the same time both the spatial and spectral resolutions improved in order of size. The highest improvement is to be waited in the field of practical data supply of radar technology. In 2002 in research-level the hyperspectral measuring technology also appeared in Hungary [23].

Biological resources

Bio-diversity and homogenous monoculture of an agricultural crop seems to be an irresolvable contradiction. The strongly fragmentised ecological islands poor in species increase the energy need of the whole agricultural region. In Hungary the spatial mosaic-like structure of soils is confirmed by the land use databases originated from the CORINE 1 : 50 000 and 1 : 10 000-scale air photos. But complete interpretation of data source will be the future task of the agro-ecological modelling. According to Horváth [17], by a spatial extension of the soil–plant–weather–pest system that can be considered homogeneous within a small plot a space specific complex ecological model can be obtained [18, 25, 26]. The models of the ecological system are strongly connected to and dependent from each other, however, each of them works by itself.

The ecological examination of spatial processes is very important in life-systems of most species. Population ecology is a branch of ecology that studies the structure and dynamics of populations. Populations can be defined at various spatial scales. Local populations can occupy very small habitat patches like a puddle. A set of local populations connected by dispersing individuals is called a metapopulation. Populations can be considered at a scale of regions, islands, continents or seas. Even the entire species can be viewed as a population [3]. Populations differ in their stability. Some of them are stable for thousands of years. Other populations persist only because of continuous immigration from other areas [10].

Sharov [37] introduced the definition of physiological time via population ecological research to describe biological maturity and time relationships. Spatial processes of ecological analysis are very important in life-systems of most of the species. They may so significantly modify system behaviour that local models would be unable to predict population changes. Several methods are used for description of the spatial processes such as random walk, diffusion model, dispersal mechanism, metapopulation analysis etc.

Random walk is simulated after several time steps until the distribution of organisms becomes close to the 1 or 2 dimensional normal distribution. The diffusion models can be applied to any initial distribution of organisms. The combination of long- and short-distance dispersal mechanisms is known as stratified dispersal. Metapopulation is a set of local populations connected by migrating individuals.

One of the main strengths of the geo-statistical analyses is the investigation of spatial variance and correlation.

Sampling and classification

With geo-statistics, the GIS analyst gains a wide range of tools to detect and describe expressions of spatial dependency in a study area through sample data sets. (Very simply, spatial dependency refers to the extent to which neighbouring points have similar attributes.) These tools contribute to an exploratory analysis of data by helping to describe the nature of spatial dependency in the study area. These descriptions may then be used to build predictive models for full surfaces. Any geo-statistical project begins, prior to sampling, with obtaining as much knowledge as possible about the distribution characteristics of the phenomenon under study. In cases where one does not have direct control over the production of sample data, the project begins by gathering ancillary information about the study area, the sampling methods, and the sampling scheme. Next, if a geo-statistical analysis is to be fruitful, it is necessary to examine the spatial arrangement of data samples visually and produce summary statistics that reveal characteristics of the sample data distribution. Detecting and interpreting special features, characteristics, or abnormalities of the data set are the first steps of exploratory data analysis, the success of which will influence subsequent interpretations of geo-statistical measures of variability and continuity. In addition to displaying a map of the sample locations with different palettes, one can analyse histograms of the attributes and obtain a statistical summary of the data. With these results in hand, better interpretations of spatial structure are likely as one begins geo-statistical analysis [13].

An exploration of the modeller (as called Spatial Dependence – IDRISI, Exploratory Spatial Data Analysis – ArcMap) which provides tools for measuring spatial variability (or its complement, continuity) in sample data. Model fitting is to build models of spatial variability with the assistance of mathematical fitting techniques.

The variogram surface is a representation of statistical space based on the variogram cloud. The variogram cloud is the mapped outcome of a process that matches each sample data point with each and every other sample data point and produces a variogram value for each resulting pair. Typically, uncovering spatial continuity is a tedious process that entails significant manipulation of the sample data and the lag and distance parameters. With the Spatial Dependence Modeller, it is possible to interactively change lag widths, the number of lags, directions, and directional tolerances, use data transformations, and select among a large collection of modelling methods for the statistical estimator. If the degree of spatial dependence decreases equally at the same rates for all sample pair separation directions, the model design is isotropic. With model fitting, the continuity structures suggested by the semi-variograms produced can be interpreted as well as any additional information we have obtained. The parameters for the structure(s) will describe the mathematical curves that constitute a model variogram. These parameters include the sill, range, and anisotropy ratio for each structure. When there is no anisotropy, the anisotropy ratio is represented mathematically as a value of 1. The sill in model fitting is an estimated semi-variance that marks where a mathematical plateau begins. The plateau represents the semi-variance at which an increase in separation distance between pairs no longer has a corresponding increase in the variability between them. Theoretically, the plateau infinitely continues showing no evidence of spatial dependence between samples at this and subsequent distances. It is the semi-variance where the range is reached.

The goal is to decide on a pattern of spatial variability for the original surface. To carry out this goal successfully with limited information requires multiple views of the

ISSN 1589 1623

variability/continuity in the data set. This will significantly increase the understanding and knowledge of the data set and the surface the set measures. Finally, in the last section of the analysis, kriging, IDW and other spatial estimators can be used, to test models for the prediction and simulation of full surfaces. There are more types of kriging, which present different results.

Ordinary kriging is known to be a Best Linear Unbiased Estimator, because it assumes that constant mean (μ) unknown, simple kriging where μ is known and universal kriging where μ(s) is some deterministic (trend) function. The final result of kriging will be to produce two images, a surface of kriged estimates and a surface of estimated variances. Kriging estimates a new attribute for each location (pixel) on the basis of a local neighbourhood. Co-kriging is another useful geo-statistical tool that uses an additional sample data set to assist in the prediction process. Co-kriging assumes that the second data set is highly correlated with the primary data set to be interpolated. Co-kriging is useful, for example, when the cost of sampling is very high and other (cheaper or available) sample data can instead [20].

The interested reader should consult Cressie [11], Isaaks and Srivastava [19] for additional explanations of kriging and general geo-statistical questions.

Data conversion or comparability

In case of field-scale data collection and processing the agricultural application of machines having satellite-base GPS means technical breakthrough. The technological decision support system building around it is called precision agriculture (PA). It has an almost decade-long past concerning from the 1992 – the year of the first American conference. Standards for agro-environmental management systems have the same short past in practice. Both areas, from the view of methodology and applicable techniques, were efficiently inspired by the quite rapid and continuously traceable development in information technology. The more stringent quality assurance and increasing agro-environmental protection requirements made the need of elaboration of complex environmental indicators obvious for analysis of environmental management and for the loading capacity analysis of the environment as a living system.

Similarly to the large-scale and efficient control functions of industrial closed systems, in open agricultural environment system analysis of material and energy flow and life cycle assessment to be shown by us become possible. The conditions of maintenance and effective operation of the environmental management systems are the continuous and reproducible measuring, tracing, improvement and development that help to increase the environmental effectiveness of production. For this purpose proper knowledge of the site and environment of the agricultural production are essential.

For the common resolution of the above-mentioned topics, an alternative is offered by the simultaneous introduction of the precision production system and the environmental management system. The precision agriculture is about to increase the effectiveness of production with optimisation of raw materials' (water, seeds, chemicals etc.) making plant growing more economic.

Determining elements of precision agriculture are: high-precision, continuous site determination, geographical information and remote sensing tools of analysis and the highly automated fieldwork [46].

This information system of precision agriculture can be the basis of establishment of the environmental management systems. Its first step is the survey and the evaluation of

ISSN 1589 1623

the environmentally active factors of the given production process and their effects. Knowing them the two most important fundamental principles, the continuous measuring–monitoring and the in parallel made continuous improvement–development can be realised.

For establishment of the environmental management systems the life cycle assessment provides help that determines the materials and energies used during the production as well as quantity and quality of the released potential toxic materials and wastes within the investigated production system. On the basis of these results the environmental effects of production can be estimated. The ISO 14040-49 family of standards contains the standards of life cycle assessment [8]. According to the examination standards, the life cycle (Hungarian Standard ISO 14040, 1997) is the subsequent, connecting stages of the influence system of a product, from purchasing the raw materials or from the formation of the natural resource to the reuse or waste disposal.

One of the problems of evaluation is the fact that effects of directly not measurable input materials must be calculated to directly not comparable effects.

In particular, this data integration process is often based on techniques for the management of uncertainty.

Bayesian probability theory has been proven to be sensitive to inaccuracies in the input probabilities. The Bayesian model often requires that events are independent of each other. This assumption is rarely true in real life. Neural networks and Fuzzy sets have an alternative approaches to handle spatial (mapped) and measured uncertainty. Keller [21] and Ultsch [47] give several examples of how to apply neural networks to environmental problems.

Uncertainty in any data layer will propagate through an analysis and combine with other spatial and attributive sources of error, including the uncertain relation of the data layer to the final decision set. In traditional GIS analysis, uncertainty is not taken into account in the database. As a result, hard decisions are made with very little concept of the risk involved in such decisions. Mays et. al. [31] demonstrate how simple it can be to work with measurement error and its propagation in the decision rule. The task of the decision maker is to evaluate a soft probability map and set an acceptable level of risk with which the decision maker is comfortable. By knowing the quality of the data, the decision maker can view the decision risk occurring across an entire surface, and make judgements and choices about that risk. Finally, any further analysis or simulation modelling of impacts with such data increases the precision of those decisions as well.

Results of the fuzzy analysis are the fuzzy layers, which are suitable for describing the decision on spatial uncertainties. Fuzzy set theory [22, 53, 54] is a mathematical method to characterise and quantify uncertainty and imprecision in data and functional relationships. Fuzziness represents situations where membership in sets cannot be defined on a yes/no basis since the boundaries of sets are vague. The fuzzy sets are classifications of data in which the boundary between classes is not distinct. The basics of fuzzy set theory have been presented and discussed in numerous articles [6, 32, 43]. Fuzzy set membership functions describe the degree to which data belongs to for each class (fuzzy set). These functions take on values between 0 and 1, and depict the grade of membership (also known as the possibility) that a certain entity has in that class. For the description of these continuous risk levels the fuzzy functions are more appropriate than the Boolean layers. Transient values correspond to risk levels in accordance with the applied functions.

ISSN 1589 1623

Data management

Data management of environmental models comparing to the business applications contains some special problems.

The amount of data to be processed is extremely large particularly due to the exponential increase of spatial and spectral resolution values of air photos and space images.

The data managing hardware and software environments are heterogeneous and divided since data owners produce their data in different environments.

Standards of new generations of digital data collectors, for example mobile phones (MMS, GPRS) have not had always-compatible data form.

In case of precision agriculture, data formats of machine computers can only partially standardised.

Environmental data objects are often spatio-temporal and frequently uncertain. For natural resources no standard object library was made that could follow behaviour of complex logical connections, nested and joint objects.

In those situations, where one works with big volume or rapidly changing data, like in ecology, it seems favourable to establish a divided database, where each database is to be installed to those users, who use them most frequently. Successful application of divided database must meet the following requirements:

- Local autonomy, namely each user place has authority over the database and the operational system.
- Thus all user places are equal.
- Conditions of continuous operation are given, even if different user places make different interventions in the system at the same time.
- Physical and logical data independence must be assured.
- Physical fragmentation of the logical field must not be perceptible for the users.
- Combined application not to cause data duplication.
- Relation database should make the data distribution possible of those answers to the different places, but to give full value data as well.
- False changes are to be deleted.
- Divided DBMS should run on different type of computers as well, so not to be dependent on hardware.
- It must operate compatible with different operation systems since it may be supposed that in each place of the organisation different operational systems had been installed.
- Due to the heterogeneous network possibility mentioned in the introduction part, the network independence is a requirement as well.
- The divided relation database must be independent from the probably different relation databases used in the user places.
- It must be safe in data protection and loss of data points of view. In case of a divided relation database risk of deterioration of the database is smaller than of central databases.

The most common used relation database is a logically planned, traditionally two-dimensional data management system. Its disadvantage is that it describes the real world in tables. Records of these tables atomise the connections and fragment them along some kind of logical organization. It is not able to reflect the embedment of the objects.

ISSN 1589 1623

Formal differences of records in columns are not able to emerge. Logical normalisations and physical planning could take a long time in complex relations. Time for reaching the data in big tables can increase.

A part of data structures making spatial indexing possible, for example the KD-tree, the 4-tree, the R+ tree, the BSP-tree are not easily matched into an object-oriented system since they cut the objects into pieces. It contradicts the essence of object-oriented planning that uses complex units [27].

Since the R-tree, the Rective-tree, the KD2B-tree and the Sphere-tree data systems handle the objects as units, they can be used properly in object-oriented modelling [33]. It makes keeping the relation systems possible without breaking the inner connections. Most of the objects in the real world have 1 : M-type relation system. This can be modelled in a higher level of abstraction. Instead of attachment of simple attributes given by the relation database, it makes input of complex attributes possible. The classification builds on the inner connection instead of the type of entity. Origin of objects in the attributes turns traceable. The attribute and the geographical data are not divided. Application of object-oriented models in description and analysis of ecological relations is expected to bring large progress.

Model creation

Prior to actual modelling certain preliminary works must be performed whose aims are to evaluate the resulting situation on the basis of the available data. Based on the calculations made by the preliminary model the optimal execution of the field exploration work, place and method of sample taking as well as planning of perception data types and perceptual frequency must be determined.

Model creation begins with preparation of a conceptual model. In formation of the conceptual model that fundamental principle becomes embodied that generally aim of the model is not to describe the process with detailed mathematical methods, but to reproduce our idea about the problem to be solved, of course in a properly controlled form. In case of the models available today the main question of model creation is not the mathematical description anymore but choosing from amongst the processes to be taken into account. Choosing the proper software the mathematical model in accordance with our decision is available.

Modelling helps to handle the complex problems, but determined by the modeller interactions that one cannot follow in head or by simple calculation methods. In case of numerical resolution methods boundaries of the resolution do not preclude any partial processes, therefore in case of uncertainty one can decide to take the process into account or neglect it on the basis of the calculations made by just the method. The system-like approach of modelling provides that one must take a clear stand on each of the parameters in the mathematical description of the processes. Neglecting a parameter is possible only with changing in the structure of the model.

Most field ecologists are not good at abstraction. If they build a model they often try to incorporate every detail. Most mathematicians are not good at interpretation of their models. Usually they think of “clean models and dirty reality”. However, both abstraction and interpretation are necessary for successful modelling. Many system properties are not represented in the model and some model properties cannot be found in real systems [37]. During the creation of a model the following points of view should be taken into account:

ISSN 1589 1623

(1) select the optimal level of complexity, but do not try to make a universal model,
(2) plan model development for each time period and area,
(3) if possible, incorporate already existing simulation or stochastic-deterministic models.

In case of ecological models three basic conditions must be cleared. These are as follows:

- type of the object (flora, fauna, etc.) and the volume of the impact to the environment must be determined,
- the medium, in which the processes take part must be determined in space and time,
- situation and response reactions of potential impact receivers must be determined in space and time.

Choosing the software is performed on the basis of choosing conceptional model, and even in this case it would be practical to set up the conceptional model independently from the software, if the available computer program is given.

In the second phase the possible simplifications, proper for the limitations of the software, must be decided on. If the conceptional model defers significantly from the modelling possibilities of the available software, it is not worth starting the detailed elaboration of the model. In case of the single model systems it is advantageous, if the user have scope for action from that point of view how the given software is able to calibrate or validate the environment.

First phase of elaboration of a model is the verification of the software. During verification we examine whether the model gives a properly exact solution in case of the known, and mainly analytically resolvable tasks. Parameters can be independent from time or changeable in time. By spatial extension they can be point-like, linear or spatial-type ones.

During the calibration of the parameters the errors of the parameters are corrected. Aim of calibration is generally not to decrease the deviations to minimum at the measuring point, but to display the character of the process as precisely as the processes taken into account according to the conceptional model and the supposed homogeneity during the parameter estimation make it possible.

Validation means the control of the calibrated model with use of such relevant cases that were not used during calibration: for example, extreme phenomena or simulation of scenario or period left out from calibration. If the results of validation are not satisfactory, the parameters must be modified or even the conceptional model too.

Boundary conditions, characteristics of the medium and the object can be changed according to the chosen scenario. These data are the basis of the evaluation, whose spatial visualization can be performed by geographical information systems.

During post-control material and energy movement as well as the effect of intervention must be traced by establishment of a monitoring network. Data received during the operation of the monitoring must be regularly evaluated, generally by with the help of modelling as well, where the expectable trends can be foretold. Estimation of the parameters of the model must be checked, and if it is not sufficient, the processes taken into account – i.e. the conceptional part of modelling – must be reviewed.

REFERENCES

[1] Andrén, O., Rajkai, K. & Kätterer, T. (1993): Water and temperature dynamics in a clay soil under winter wheat: influence on straw decomposition and N immobilization. – Biol. Fertil. Soils 15: 1–8.
[2] Antal, E. (1978): Éghajlati erőforrás és mezőgazdaság jelentősége. [Climatic potential and and importance of agriculture.] – Időjárás 76: 602–313. Aszály 1983. [Drought 1983.] KÉE, Budapest, pp. 9–43.
[3] Berryman, A.A. (1981): Population systems: a general introduction. Plenum Press, New York.
[4] Bíró, T. & Thyll, Sz. (1999): A belvíz-veszélyeztetettség térképezése. [Risk mapping of excess water.] – Vízügyi Közlemények 81(4): 709–718.
[5] Burgos, J.J. (1986): Equilibrium and extreme climatic conditions of world's biomass and agrosystems. Land use and agrosystem management under servere climatic conditions. – WMO Technical Note 148: 12–56.
[6] Burrough, A.P. & McDonnel, A.R. (1998): Principles of Geographical Information Systems. Oxford University Press, Oxford, 327 pp.
[7] Cambell, I.M. (1977): Energy and atmosphere. A physical–chemical approach. John Wiley and Sons Ltd, London.
[8] Cascio, J., Woodside, G. & Mitchell, P. (1996): ISO 14000 Guide, The New International Enviromental Standars. McGraw-Hill, New York, pp. 3–217.
[9] Csáki, Cs., Harnos, Zs. & Láng, I. (1984): Agricultural development and ecological potential: the case of Hungary. Kieler Wissenschaftsverlag Vauk
[10] Clark, L.R., Geier P.W., Hughes, R.D. & Morris, R.F. (1967): The ecology of insect populations. Methuen, London.
[11] Cressie, N.A.C. (1993): Statistics for spatial data. Wiley and Sons Publisher, New York, 887 pp.
[12] Csáki Cs., Harnos, Zs., Rajkai, K. & Vályi, I. (1988): Hungarian agriculture: development potential and environment. – In: Parikh, J.K. (ed.): Sustainable development in agriculture. IIASA, Martinus Nijhoff Publishers, Dordrecht, The Nederlands, pp. 253–295.
[13] Eastman, J.R. (1999): Idrisi32 tutorial. Worcester, MA USA, 291 pp.
[14] Environmental Management Standard. McGraw-Hill, New York, pp. 3–81.
[15] Harnos, Zs. (1991): Az alkalmazkodó mezőgazdaság rendszere. [Systems of sustainable agriculture]. – Módszertani kutatások, Kertészeti és Élelmiszeripari Egyetem, Budapest, 1–91.
[16] Harnos, Zs. (1993): Időjárás és időjárás-termés összefüggéseinek idősoros elemzése. [Time series analysis of climate-yield relationships.] – In: Baráth, Cs., Győrffy, B., Harnos, Zs.: Aszály 1983. [Drought 1983.] OTKA kiadvány, pp. 9–43.
[17] Horváth, L. (2002): Térbeli inhomogenitások kezelésének módszertani problémái a precíziós növénytermesztésben. [Spatial inhomogenity problems in precision farming.] – In: Proc. Agrárinformatika 2002, Debrecen, pp. 335–343.
[18] Hufnagel, L. & Gaál, M. (2002): Többváltozós állapotsíkrendszerek alkalmazása valós és szimulált adatsorok kezelésében. [Multivariate state-planes in service of analysis of field- and simulated data.] – In: Proc. Agrárinformatika 2002, Debrecen, pp. 315–325.
[19] Isaaks, H.E. & Srivastava, M.R. (1989): An introduction to applied geostatistics. Oxford University Press, Oxford, pp. 3–553.
[20] Johnston, K., ver Hoef, M.J., Krivoruchko, K. & Lucas, N. (2001): Using ArcGIS Geostatistical Analyst. ESRI, Redland, USA, 287 pp.
[21] Keller, H.B. (1995): Neural nets in environmental applications. – In: Avouris, N.M. & Page, B. (eds.): Enviromental Informatics – Methodology and Applications of Environmental Information Processing. Kluwer Academic Publisher, Norwell, Mass., pp. 142–165.
[22] Klir, G.J. & Yuan, B. (1995): Fuzzy Sets and Fuzzy Logic. Prentice Hall, New Yersey, 150 pp.

ISSN 1589 1623

[23] Kardeván, P., Vekerdy, Z., Róth, L., Sommer, S., Kemper, T., Jordán, Gy., Tamás, J., Pechmann, I., Kovács, E., Hargitai, H. & László, F. (2003): Outline of scientific aims and data processing status of the first Hungarian hyperspectral data acquisition flight campaign, HYESNS 2002 HUNGARY. – In: Proc. 3rd EARSeL Workshop on Imaging Spectroscopy, 13–16 May 2003, German Remote Sensing Data Center, Oberpfaffenhofen, Germany (*in press*).
[24] Kovács G.J., Ritchie J.T. & Németh, T. (1995): Testing simulation models for assessment of crop production and nitrate leaching in Hungary. – In: Agricultural Systems, Vol. 49., No 4. Elsevier Sci. Ltd., Oxford, England, pp. 385–397.
[25] Kovács, G.J. & Nagy, J. (1997): Test runs of CERES-maize for yield and water use estimation. In: Nagy, J. (ed.): Soil, plant and environment relationships. Current plant and soil science in agriculture, Debrecen, pp. 115–119.
[26] Ladányi, M., Horváth, L., Gaál, M. & Hufnagel, L. (2003): An agro-ecological simulation model system. – Applied Ecology and Environmental Research 1(1–2): 47–74.
[27] Laurini, R. & Thomson, D. (1992): Fundamentals of spatial information systems. Academic Press, London, 673 pp.
[28] Láng, I. (ed.) (1983): Mezőgazdaság agroökológiai potenciálja az ezredfordulón. [Agro-ecological potential of agriculture on millennium.] Tárcaközi Bizottság Jelentése, Budapest.
[29] Láng, I. (2002): Agrártermelés és globális környezetvédelem. [Agricultural production and global environmental protection.] Mezőgazda, Budapest, 211 pp.
[30] Láng, I., Csete, L. & Harnos, Zs. (1988): Resources Adjustment and Farming Structures. – European Review of Agricultural Economics 15: 2–3.
[31] Mays, M.D., Holzhey, C.S., Bogardi, I. & Bardossy, A. (1995): Managing risk and variability with fuzzy soil interpretations. – In: Robert, P.C., Rust, R.H. & Larson, W.E. (eds.): Proceedings of site-specific management for agricultural systems. American Society of Agronomy, Madison, WI, USA, pp. 187–200.
[32] Moore, I.D. (1996): Hydrological Modeling and GIS. – In: Goodchild, M.F., Parks, B.O. & Johston, D.C. (eds.): GIS and Environmental Modelling. Progress and Research Issues.
[33] Oostrerom, P. (1993): Reactive data structures for geographical information systems, spatial information systems. Oxfords University Press, 198 pp.
[34] Pálfai, I. (1994): Az Alföld belvíz-veszélyeztetettségi térképe. – Vízügyi Közlemények 76(3): 278–290.
[35] Racskó, P. (1991): Some aspects of crop modelling. Workshop on Agricultural modelling, Visegrád, 25 pp.
[36] Swaminathian, M.S. (1991): From Stockholm to Rio de Janeiro. The road to sustainable agriculture. M. S. Swaminathan Research Foundation Madras, 68 pp.
[37] Sharov, A.A. (1992): Life-system approach: a system paradigm in population ecology. – Oikos 63: 485–494.
[38] Somlyódi, L. & Straten, G. van (eds.) (1986): Modelling and Mapping shallow lake eutrophication – with application to Lake Balaton. Springer Verlag, London.
[39] Somlyódi, L. (2000): Magyarország vízgazdálkodási stratégiai kérdései. MTA Budapest.
[40] Somogyi, S. (ed.) (2000): A XIX századi folyószabályozások és ármentesítések földrajzi és ökológiai hatásai. [Geographical and Ecological Impacts of River Regulation in XIX century.] MTA Geographical Science of Hungarian Academy, Budapest, 302 pp.
[41] Stefanovits, P., Filep, Gy. & Füleky, Gy. (1999): Talajtan. [Soil Science.] Mezőgazda, Budapest, pp. 15–470.
[42] Szász, G. (1985): A klímapotenciál fogalma és alkalmazása a mezőgazdasági termelésben. [Definition of Climapotential and Adaptation in Agricultural Production.] – In: A klímapotenciál és az agrometeorológiai információk népgazdasági hasznosítása. Magyar Meteorológiai Társaság, Budapest.
[43] Tamás, J., Csillag, J. & Murányi, A. (1998): Risk mapping of potential heavy metal pollution modelling combined effects of pH, clay mineral and organic matter. – In: Filep, Gy. (ed.): Soil pollution. Soil, water and environment relationships, Debrecen, pp. 225–239.

ISSN 1589 1623

[44] Tamás, J. & Nagy, J. (1997): Evaluation of irrigation schedules by CROPWAT FAO model. Soil, plant and environment relationships. – In: Nagy, J. (ed.): Current plant and soil science in agriculture, Debrecen, pp. 115–119.
[45] Tamás, J. & Bíró, T. (2001): Vízkészlet modellezés. [Modelling of water resources.] University of Debrecen, 201 pp.
[46] Tamás, J. (2001): Precíziós mezőgazdaság. Mezőgazdasági Szaktudás Kiadó, Budapest, 164 pp.
[47] Ultsch, A. (1995): Einsatzmöglichkeiten von neurolanen Netzen im Umweltbereich. – In Page, B. & Hilty, M.L. (eds.): Umwetinformatik- Infortikmethoden für Umweltforschung. Oldenbourg, Munich, Vienna.
[48] Varga-Haszonits, Z. (1981): A gazdasági növények terméshozamának éghajlati potenciálja. [Climatic potential of yield of the crops.] – MTA Közlemények 14(2–4): 253–270.
[49] Varga-Haszonits, Z., Varga, Z., Lantos, Zs., Vámos, O. & Schmidt, R. (2000): Magyarország éghajlati erőforrásainak agroklimatológiai elemzése. [Agroclimatic Analysis of Hungarian Climatic Potential.] Mosonmagyaróvár, 219 pp.
[50] Várallyay, Gy. (1985): Magyarország 1:100 000 méretarányú agrotopográfiai térképe. – Agrokémia és Talajtan 34: 234–248.
[51] Várallyay, Gy. (1993): Soil data-bases for sustainable land use: Hungarian case study. Soil resilience and sustainable land use. – In: Greenland, D.J. & Szabolcs, I. (eds.): CAB International, pp. 469–495.
[52] Webster, R. (1977): Quantitative and numerical methods in soil classification and survey. Clarendon Press, Oxford.
[53] Zadeh, L. (1965): Information and Control 8**:** 338–353.
[54] Zimmerman, H.J. (1985): Fuzzy set theory and its application. Martinus Nijhoff, Dordrecht, 363 pp.

ISSN 1589 1623