Sustainable Development of Rural Communities in Poland – an Attempt to Apply Genetic Algorithms and Expert Systems in Decision Making Process

Radosław Bułat¹, Łukasz Popławski²

¹Faculty of Philosophy, Jesuit University Ignatianum in Cracow, Cracow, Poland
²Faculty of Finance and Law, Cracow University of Economics, Cracow, Poland
E-mails: ¹radoslaw.bulat@ignatianum.edu.pl (corresponding author); ²lukasz.poplawski@uek.krakow.pl

Received 11 February 2017; accepted 03 April 2017

Abstract. All scientific projects have the need of operating on large matrices of data. This article has the aim to establish if there is a correct and useful scientific method of data management by utilizing genetic computations and/or artificial intelligence expert systems in the sustainable development of selected rural areas in Poland, using a cross-field project of economic development and IT. It discusses the methods of preparing of the data from their most basic form, data manipulation, database operations in the form of genetic operators, custom scripting and working algorithms, getting the answers from the multi-criterion analysis. There are also further case studies, achievable with the methodology, as a successful example of the right set of methodologically correct tools. Also a consideration if there is a potential for its possible usage on economic and social analysis in the aspect of sustainable development.

Keywords: sustainable development, resource management, genetic algorithms, expert systems, business intelligence, optimization problem.

JEL Classification: C61, Q01.

Conference topic: Informational Society: Challenges, Solutions and Tendencies.

Introduction

As is known, economical calculations on a large scale needs to operate on enormous amount of data for analytical and statistical purposes. In most cases, there are demographic or financial surveys from various regions and entities, which need to be processed and calculated. The next course of action is logically making a prediction or assumption about future trends or calculating a min-max solution, which would utilize all known resources in the most effective manner, often as a balancing act of increasing the well being of some regions more than the others. (Ginevičius, Brauers 2011)

In this paper a solution based on genetic algorithms and high level programming language will be presented, to test such an approach on a real world scenario and data, on the attempt in aiding the sustainable development of rural communities in Swietokrzyskie Voivodeship in southern Poland, by optimizing their territorial budget allocation. There has been a lot of surveying in those communities over the past years, and their data is a good example of statistical sample. There’s also the urgent need for the local governmental organization to have access to all the correlates the data can provide, thus giving the need to use the genetic application outside the lab environment by local executives (Song et al. 2016). The main aim of this article is to present an example of applying genetic algorithms and expert systems in decision making process on a sustainable development of rural communities in Poland as a viable way of assisting those communities in their decision-making processes.

About the sustainable development

According to (Rutkowska-Podołowska, Popławski 2014) – the first-ever definition of sustainable development was provided in the 1987 report of the UN World Commission on Environment and Development: Our Common Future. The Commission Head was then Ms Gro Harlem Brundtland. The Report defined sustainable development as “development that meets the needs of the present without compromising the ability of future generations to meet their own needs”. On the other hand, Article 3 of the Polish Law on Environment Protection of 27 Apr 2001 (Polish Government 2001) states: “sustainable development is a socio-economic development which integrates political, economic and
social activities while preserving the ecosystem balance and continuity of basic natural processes so as to guarantee meeting the basic needs of individual communities and individuals of the present and future generations”. Again, Article 5 of the Constitution of the Republic of Poland (Polish Government 1997) provides that the state “assures protection of the environment, observing the principle of sustainable development”.

“The sustainable development is thus a doctrine of political economy, which sets up the quality of life on the level, which is afforded by the contemporary civilization development. The essential and quite often neglected message is included in the beginning of a famous sentence from the WECD Report from 1987, stating that: “at the present civilisation level the sustainable development is possible, which means such a development, in which the needs of contemporary society can be met without decreasing chances of the future generations for their fulfilment”. The authors of the Report stressed, that civilisation reached a prosperity level, which is possible to maintain, provided that it is managed properly, i.e. change of the economic growth quality, which should be performed by linking economic interests with environment protection. A model of such an economy presumes appropriately and consciously created relations between the economic growth, environmental care (not only natural environment but also the artificial one – created by a man) and human health. The doctrine of a sustainable development aims at social justice, by enterprises economic and environmental effectiveness. In accordance with the sustainable development a society should live, as much as possible, from the ‘interest’, and not from the ‘capital’ (Rutkowska, Rembielak-Vitchev 2008). The sustainable development means, that an economic growth leads to increasing the social integrity (decrease in social stratification, equalising the chances, preventing from marginalisation and discrimination) and increasing the quality of the natural environment by limiting the harmful influence of production and consumption on the environment, and environmental resources protection (Siekierski, Rutkowska 2008; Lankauskienė, Tvaronavičienė 2013).

The local situation

The area covered by protected areas in Poland amounts to 33.1%; the region with the highest percentage of areas covered by them is in the Świętokrzyskie Voivodeship – 62%. There are 72 municipalities (communes) in the Świętokrzyskie Voivodeship among which 49 protected areas cover more than the half of the Świętokrzyskie area. These areas are usually landscape parks and protected landscape areas.

The protected areas are in most cases situated in rural areas, in only few cases they are situated in the areas of the cities e.g. Szczecinski Landscape Park. The location of the protected areas in the rural premises gives a reason to consider both functions of these areas as a whole. It is even more necessary in the case of landscape parks, because of the fact that the agricultural use, according to regulations, is allowed there. Living conditions of the protected areas and landscape parks’ inhabitants are determined by the legal restraints (Popławski 2005).

Landscape parks as protected areas shall according to the regulations protect the environment on its territory, while development of rural areas is mostly closely connected with agriculture. Those two elements sometimes cause conflicts between the farmers and park management. That is why the courses of development of protected areas are important arguments for starting the researches on development possibilities of those territories (Lojewski 1998). The findings of the researches on rural areas’ development possibilities in accordance with environment protection and considering all the aspects of development, might resolve the conflicts or at least ease them. The interest of all the people in a given protected area should be taken into account and pointed out while still obeying the regulations presently in force.

The direct courses of development depend on specific conditions in every commune.

Formation of development strategy

Planning sustainable development (eco-development) of protected areas in communities should lead through legal, economic and social actions, and especially spatial ones to:

– the compliance of a character and a structure of development, including a level and intensity of development as well as spatial management with features and values of the environment of a given natural area,

– exposing natural, cultural values and landscape ones of the environment and their harmony with spatial management together with their protection,

– creating spatial conditions providing protection of unique values and merits of natural and anthropogenic environment together with a chance to regain the lost ecological balance through suitable actions, including spatial ones as they are common goods on a transnational scale.

Such an action aims at accepting a concept of development by society in order to achieve social and economic prosperity without a simultaneous threat to natural resources. The acceptance is a condition to transport the economy of the region in compliance with the programme of eco-development. Therefore, among a lot of actions undertaken within policy of eco-development, it is necessary to enhance social awareness about an individual responsibility of each citizen for implementing environmental protection and protection if its resources in a workplace, a place of living and rest. Efforts should be made to make each human being more active, responsible and better prepared for the reali-
zation of the requirements which include maintaining the quality of the environment and the quality of life and providing sustainable development. Elaboration of the realization of sustainable development should be based on programming it on the basis of possessed information and prepared variants considered and assessed from the point of view of the commune mission. The programme of sustainable development will be the document realizing the programming for this type of integrated actions consisting of (Wiatrak 2003):

- the strategy of sustainable development for the commune,
- the policy of land management of the commune,
- the ecological policy of the commune (the literature on the subject also enumerates interchangeably: the programme of environmental protection).

The implementation of the concept of sustainable development, to a large extent, depends on an activity of local governments, including active participation of a society because the spatial structure of protected areas use results from the activity of regional environments and entities, expressed in concrete projects and undertakings based on taking three orders (economic, social and ecological) into account while realizing them.

It should be emphasized that IT tools to undertake certain actions presented below can be used on all stages of implementation of the concept of sustainable development by the authorities of the commune.

The comprehensive planning system has to be of an integrated nature through parallel actions in three strategic spheres (Markowski 2011):

- economic (commercial),
- social,
- institutional and organizational as well as procedural (procedures and a range of arrangements of planning documents).

The local authorities are responsible for development in numerous areas in the territory of their influence through creating the conditions for economic development. Creating the conditions is based on a lot of actions, which especially include: land use and development of infrastructure, development of education and creating the conditions to support intellectual and social capital, creating an offer for initiatives for enterprising attitudes and administrative services as well as supporting investment actions for national and international capital in the scope of investment. Furthermore, in the financial scope, supporting any investment activity especially in the scope of development of new technologies and industry through, among others, ordering to carry out appropriate investments or providing institutional aid (renting premises, making business easier – granting approvals) and tax allowances in the scope of their power.

Owing to their legislative competences, the local authorities should make use of financial, administrative, strategic planning and operational instruments and frequently as a dominant subject (the biggest employer) in order to support development in accordance with the concept of sustainable development. Currently, the most important problem of implementing this concept in Poland is bringing the economic and social interests with the ecological aims, and especially with the change of the carbon economy into the economy based on the mixture (carbon or other fossils plus various types of energy from renewable resources). Implementing the concept of sustainable development is especially crucial in the protected areas on account on a necessity to combine preservation of nature with the economic and social goals (the primary aim protection of these goods against the increase in well-being or prosperity and inhabitants’ aspirations). To remedy these problems of local development in protected areas, it is necessary not only to implement a strategy of local development or local development plans, but an integrated programme taking plans of protection of nature of a given protected area and programmes of environmental protection into consideration.

There are a lot of arguments in favour of active support of actions for protection of the discussed areas and appropriate management, but first of all they are the following from the economic and social point of view:

- the so-called external costs can be limited, the quality of production improved, the resource-intensive economy decreased, the losses caused by environmental pollution minimized, the working, accommodation and rest conditions improved in the protected areas through the limitation of the economic actions less friendly to the environment or simply detrimental to it, even if it means to tackle the incompatibility of economic aims (Tvaronavičienė et al. 2015).
- financial resources for environmental protection increase a demand for devices for its protection, which stimulates development of industrial sectors producing protective devices as well as decrease expenses for health care, resulting from ecological reasons;
- these areas contribute to maintenance and protection of the natural and cultural heritage, and at the same time increase inhabitants’ ecological awareness and historical identity;
- owing to environmental protection, new workplaces are created, for instance, more employees are needed on ecological farms as well as in the service in the field of environmental protection.

 Nonetheless, even the best elaborated concept of eco-development or sustainable development will not have desired effects if it is not implemented and realized. It requires undertaking a lot of various actions on all levels of governance, using instruments and tools helpful in its realization, and also persuading people about its reasonableness, effectiveness and profitability.
The aggravation of inequality in the scope of the pace and level of social and economic development refers to numerous areas, including those precious natural ones, which requires defining priority goals and implementing a strategy which could enable to level developmental chances, facing the existing disproportions. Monitoring and a control of the applied strategy realization are necessary to solve them effectively. Since efficiency of the system depends on the condition and quality of the existing potential, including institutional infrastructure, which will be controlled in the interest of a local community. Awareness among the activists of economic life is important that possessed resources and undertaken decisions will determine the quality of the local environment in the future.

Business activity which should be promoted in protected areas can be divided into several types, especially pro-ecological agriculture together with food processing and rural tourism with services and trade. The development of tourism (ecological tourism) ought to be supported and treated as an alternative source of income for farmsteads. A similar point of view presented (Gaule, Zilinskas 2013; Grzelak 2008; Biernat-Jarka 2010; Jarka 2012; Pilvere 2012). Active environmental protection also belongs to them within cooperation with the administration of protected areas, especially the management of landscape parks.

Data format and preparation

In the most common case, the data about those communities come in small chunks from local surveys, often done by hand. The digitization of those data is a must, and most often need either a skilled typist (older method) or an OCR (image recognition) software. In both cases data from those standardized surveys come into the input of computer system, in standardized format – the most common used method are Excel (or any compatible brand) spreadsheet. Since the format also has a big impact on the processing speed and overall computing power need, the spreadsheet is not the medium which can operate on very large and complex amounts of data, even if it can cope with simple analysis and statistical metrics (Popławski, Bułat 2008).

The solution seems to be a database system, the most commonly used would be an SQL Server (for commercial work – it’s licensed software) or a MySQL server (for internal and non-profit work, as it is a free solution). An SQL server (an abbreviation from Structured Query Language) allows for greater storage, flexibility and programmability of the data structures. After the importing process (which encompasses data structuring, and transposing between spreadsheet, comma separated values and in some cases XML files (eXtensible Markup Language)) the database is filled with records of uniform structure, without regard of the source and method of collecting the data. (Popławski, Bułat 2008). The uniformity is a first step to getting an accurate measurement across the whole scope of measured quantities and processes.

Scripting and querying the expert system

Having all the data neatly structured and packaged, there’s the need of obtaining BI (Business Intelligence – a process for extracting useful information from raw data). There’s the need for building an expert system, translating the query into the digital form, which could be understood and executed by database system. On the inner level the best practice is using the SQL language, on the outer – some scripting, object-oriented language (such as Python or Ruby). (Popławski et al. 2011).

The first step in this method is to plan, what kind of questions the machine should be able to answer and in what form would the numeric answer be represented, what at are changeable variables in all of them and which should be dictated by users (higher tier executives and managers or business analysts). Next, those can be translated into the set of logic rules – for example a variable or a set of co-variables over or under a certain threshold, always have implied an event that is a consequence for them, and may be the beginning or end of a trend. Additionally as the classic logic only operates on a true or false model, the fuzzy logic has to be introduced which can operate on the more fluid probability scale. Again, analysing the data and history patterns in dataset, one could infer that certain trends have happened in a set percentage of cases, and thus, the results can also include a risk assessment in the form of natural language response “the predictions based on past observations can calculate X% of risk for an investment with asked parameters” – which would be an invaluable tool for a non-technical investment managing personell. (Kisielewicz 2011).

When this most difficult task of gathering all of those trends (from the general knowledge, scientific research, known good practices and other sources) is done, there exists a Bayesian Network mapping all, or most of, connections between the causes and their effects. More importantly, it allows not only a forward chaining logic reasoning (inferring an effect from the causes), but also even more powerful and useful backward chaining – which might prove invaluable in the research field. Backward chaining allows reasoning about the causes of a known effect, which might allow to see previously unknown relations between datasets – again presented with their probability values and left for testing as probable hypotheses or open avenues for further research.
Genetic computations and algorithms

Another problem encountered when working on data sets, apart from perceiving trends, is an optimization solution. There are often occurrences of real business and budget planning practices, which would need a min-max solution or achieving the highest possible payout with all the materials or resources available (for example Ginevičius, Brauers 2013). The optimal solutions are most often single peaks of complex functions in a mathematical or statistical models and as such are hard to be obtained via traditional queries of expert systems – there are simply too many variables to be accounted for, to have the feasibility of calculating the optimal solution to a given function. The analytical or algebraic methods are proven to work and find the correct solution, but the more complex the problem is, the more time and computational power it requires, and the answer that would be obtained in a matter of months or years is worthless from an economic point of view. Most often, a slightly suboptimal solution that can be done in matter of seconds, far outweighs the usefulness of the perfect one with a longer processing time – a local function extremum on a broad enough scope of possibilities is preferable to a global one.

As with all too complex arithmetic problems, there’s much usefulness in the Monte Carlo method of approximation of the function integrals and other parameters. A computation can be done remarkably faster, although at the cost of lower precision due to stochastic variations. Also the electronic random number generators permutations aren’t infinite, and as such there’s only a fixed amount of results that could be achieved in this way. Still, often the precision is enough for the application in everyday decision making process. The most creative expansion of Monte Carlo methods in modern data mining are the genetic computations. (Poplawski, Bułat 2010; Song et al. 2016).

On the algorithm side of implementing the system, there is again the need for both the correct representation of data and the criteria that should be optimised, best written in the form of complex arithmetic function with a large set of variables. The expert system that has been established in the last step can form the basics of such function – putting all of the represented and inferred data connections into a graph with node for every variable accounted for. In such a net, in theory every possible combination of variables (function arguments) can be tested and evaluated for feasibility, but as is shown before, there’s no possibility to use that route, given the physical constraints. Only a subset of all possible solutions has to be chosen for testing, specifically selected from only the solutions with the highest probability of being optimal. Thus, a genetic algorithm is enforced on the system graph, as described below:

Data representation

As have been said earlier, all data in SQL database has already been digitized and the graph including all the variables has been completed. Now, for the algorithm to work, all of the possible solutions in decision making process (e.g. branches of budget for a given rural community or places of transit for optimisation of transport orders (Bulat, Radziszewski 2011; Ardjmand et al. 2016)) have to be represented as single entities of ordered numerical values, representing every variable taken under consideration, also with noting what are all possible values every variable can get. Such data records can get very long and are analogue in their design to the DNA strands with a similar usefulness. For a starting optimisation point, a few possible solutions are generated – these can be real ones, already being used in some of the communities, but it is important to also use some of them being randomly generated – to allow the starting pool to get an uniform sample from the solutions’ space.

Data manipulation

Given the pool of fixed, random or semi-random (and thus, sometimes utterly useless) solutions, an evaluation function should be used, marking every one of them with a fitness value, signifying its usefulness in the overall process. (i.e. the projected investment return for a given budget project or a inferred route length for a logistic one). Thus all the tested solutions are ranked from the best to worst, the latter being subsequently deleted – “eliminated” from the gene pool. As the optimization is multi criterion, there could be more than one fitness value, each one emphasizing different aspect of the problem (although measured in some kind of objective uniformized units (Ginevičius, Brauers 2011)), and allowing for example to choose between cost-energy or resource-saving, as it is not often possible to maximize only one of those factors, without sacrificing the sustainability in the process. Still, even with a very large initial random search space, the remaining answers are very suboptimal in most cases. The next step in the algorithm solves this problem.

Genetic operators

There are two operations ensuring that computations will not get stuck in a local extremum of the problem space. Since all solutions with high values of fitness function tend to be grouped in an extremum, be it local or global, a mechanisms based on natural world are put in place:

– Mutation – one, or more, of elements of the already discovered most fit subset, undergo a random variation (with mechanisms ensuring that the whole solution remains feasible, even if not optimal). The randomness often makes the solution no longer fit, but sometimes can lead to breakthrough in the process, allowing to explore more of a theoretical answers space, which might contain another function extreme point.
— Crossover – one, or more, pairs of the most fit solutions are being combined, resulting in a pair of another data strands, containing a combination of both “parents” arguments strings. This process allows to keep some already calculated as advantageous clusters of data, while also allowing for experimentation in finding some others that would complement them better than the currently proposed ones.

Of course, there is no guarantee on the superiority of the “offspring” solutions (only their viability is assured by algorithmical constraints, which confine them into the subspace of solutions adhering to the baseline conditions set by system users). The “evolution”, or, more precisely the culling of the solution space leaving only the solutions with highest values of fitness function, helps to steer the algorithm towards the function limit, keeping it from reaching outward from it in next “generation” (iterations of the whole process). An example of simple algorithm using the aforementioned methods can be seen on Figure 1.
The algorithms

Having been given the methodology and its operators, there’s also a need to select the correct variant tailored to the database containing specific data – the order of operations, their count and prevalence of some of them over the others all have an impact on the overall solution quality. Over the years, there has been proven a lot of sample algorithms used in various tasks (Goldberg 1989). One of our own case study in the logistic sector has shown promising results in the application of SPEA (Strength Pareto Evolutionary Algorithm) and its variant SPEA2 (Zitzler et al. 2001). The system optimising the route planning for a pickup-and-delivery problem has proven the methodology as being correct, and from all tested possibilities, the ones mentioned above were chosen for their overall efficiency (Bułat, Radziszewski 2011). Also – similar logistics problems with multi criterion analysis were also being successfully solved recently (Jokšienė, Žvirblis 2016.)

Conclusions

To conclude, there is a high probability that in the future, the datasets on which the local government will have to work upon will increase exponentially in size as new and more variables will be collected from the increasingly environmentally-aware communities. As is proven, the methods to work on such datasets can be constructed and applied, tailoring the algorithms and control variables to the structure of each problem, aiming to make the system available for use by non-technical executive personnel of the government. There’s a possibility to construct a working system, capable of administering the expert assistance of data analysis to help with the balancing act of multi criterion optimisation. With this work, the decision making process would be possible to be delegated to the rural communities themselves without the need of central managing board – which in turn could be expected of raising the self-awareness on sustainable development in those communities as well.

These tools have not yet been distributed on a large scale due to technical and technological barriers in the introduction of the information society and the trouble of irregular covering of rural municipalities with wireless network in Poland. Work is currently underway on an increasing prevalence of e-government as a knowledge-based economy. After the implementation of e-government in the Polish governing institutions, the next step will be the introduction of some of the solutions and processes described in this work, to facilitate decision-making for low levels of regional government administration - in order to increase autonomy and decentralization of power at the local level to better meet the needs of the local communities. This proposal is a kind of novelty, not previously published, and is the early step of a broader research on its conceptual stage. Pilot use of genetic-assisted databases was conducted by the authors for 45 municipalities of Swietokrzyskie Voivodship in Poland. The resulting data from the local government units (rural municipalities) was an important element for the diagnosis of their socio-economic development.

Disclosure Statement

Both authors hereby confirm that they not have any competing financial, professional, or personal interests from other parties.

References


Popławski, L. 2005. Perspectives of multifunctional development of rural areas in the Nida Basin Landscape Park (e.g. Imielno commune), Ekonomika ir vadyba: aktualios ir perspektyvos 5: 299–305.


