LUMA-GIS Thesis nr 11

# Advanced Decision Support Methods for Solving Diffuse Water Pollution Problems

**Svajunas Plunge**

2011
Department of Earth and Ecosystem Sciences
Division of Physical Geography and Ecosystem Analysis
Centre for Geographical Information Systems
Lund University
Sölvegatan 12
S-223 62 Lund
Sweden

# Advanced Decision Support Methods for Solving Diffuse Water Pollution Problems

Svajunas Plunge
Department of Physical Geography and Ecosystem Analysis
Lund University, Sweden

Master thesis 30 credits, in partial fulfillment of the requirements
For the degree of
Master in Geographical Information Science
Supervisor: Andreas Persson
2011

LUND UNIVERSITY

# Abstract

Dealing with water diffuse pollution is a major problem for watershed managers. This problem raises many complicated questions, which are important to answer in order to reach water environment protection goals. This study suggested some possible answers for the country of Lithuania. Among them were the identification of critical source areas, the identification of sensitive areas and the application of multi-objective spatial optimization. Those decision support methods were not only suggested, but also examined through literature review and their application was demonstrated practically on the Graisupis river catchment, which is located in the middle of Lithuania. For this purpose, the SWAT (Soil and Water Assessment Tool) model was prepared and successfully calibrated and validated for water flows and nitrate load simulations. The model was calibrated for 7 years (2000-2006) and validated for 3 years period (2007-2009). The model was run for 10 years period (2000-2009) in order to obtain results for decision support methods. Critical source areas were defined as those areas, which have nitrate loads to surface water bodies higher by two standard deviations from average in the catchment. Sensitivity (nutrient leaching potential) of areas was assigned based on the response of modeled physical nature to the addition of nitrogen fertilizers. The SWAT model was also used for the simulation of effects of best environment practices. The results were imported into the genetic algorithm, which was used for the purposes of multi-objective spatial optimization. Model results indicated average nitrate loading of 15.9 kg nitrate nitrogen per hectare in the catchment. The identification of critical source areas located 12.4% of the Graisupis river catchment as risk areas. The sensitive areas identification assigned medium or low sensitivity to 99.5% of the catchment. Only 0.4% of the catchment territory was identified as high or very high sensitivity. Multi-objective spatial optimization increased the cost-effectiveness of diffuse pollution abatement 24 times (up to 50 times with lesser implementation scale), if compared to the random selection of best environmental practices. Optimization with equal weights for environmental and economic objectives resulted in 16.9 LTL for reduction of 1 kg nitrate nitrogen to surface water bodies, while providing 62% reduction of total loads to surface water bodies. This scenario required 24% of additional catchment territory to be converted to grasslands and consideration of filter strips for 34% of the catchment territory. Optimization for obtaining Pareto optimum between environmental and economic objectives provided the most cost effective solution of 9.7 LTL for reduction of 1 kg nitrate nitrogen, while providing 25% reduction of total loads to surface water bodies. This scenario required the application of cover crops on 2.6%, new grasslands on 1.6% and consideration of filter strips on 11% of the Graisupis river catchment area. Optimization for obtaining Pareto optimum between environmental and economic objectives also provided quantifiable relationship between economic and environmental objectives in the form of regression equation.

**Key words:** diffuse pollution, non-point pollution, critical source areas, sensitive areas, multi-objective spatial optimization, best environmental practices, best management practices, SWAT model, Lithuania.

# Acknowledgements

I would like to thank Dr. Ausra Smitiene for providing valuable comments and important data for the model preparation.

I would also like to thank the administration of the Lithuanian Environmental Protection Agency for allowing me to complete the model preparation part during working hours.

# List of Abbreviations

BEP[1] – Best Environmental Practice

BMP[2] – Best Management Practice

CSA – Critical Source Area

DEM – Digital Elevation Data

FAO – Food and Agriculture Organization

GA – Genetic Algorithms

GIS – Geographical Information Systems

HRU – Hydrological Response Unit

HWSD – Harmonized World Soil Database

LEPA – Lithuanian Environmental Protection Agency

LHMS – Lithuanian Hydrometeorological Service

MOSO – Multi-Objective Spatial Optimization

PEM – Point Elevation Data

PI – Phosphorus Index

SWAT – Soil and Water Assessment Tool

USLE – Universal Soil Loss Equition

WMI – Water Management Institute

1 BEP used in the European context.
2 BMP is used in the North American context.

# Table of Contents

# Preface

Ideas for this work came from the problems I encountered when working in the Lithuanian Environmental Protection Agency (LEPA). This work has been intended as a demonstration how some of those problems could be solved. Although to apply the suggested solutions on the country basis would require much more work to be done and more specialists to be involved, decision making for watershed management would reach the new level in Lithuania if these solutions were applied. This level is of sound science and methods application in decision support. Hopefully, this work will be a stepping stone to this direction.

# 1 Introduction

Diffuse water pollution is the biggest challenge for watershed managers if the deterioration of water ecosystems is going to be stopped and reversed. Current legislation such as Water Framework Directive (European Union) and Clean Water Act (USA), require substantial improvement of water quality. This is not possible without solving diffuse pollution problem. Compared to point source pollution (solved by the installation of waste water treatment plants, as well as change in chemical use and technologies), a successful solution for diffuse pollution requires much more complicated approaches. It requires the application of Best Management Practices (US terminology) or Best Environmental Practices (EU terminology) in the right complexes and in the right places. Even though a more intensive application of abatement measures would increase the possibility to reach the desired result, due to financial constraints, lesser scale of abatement is desired. In order to combine both environmental and economical goals in one assessment some kind of multi-objective functions are needed. They should also be integrated with spatial optimization techniques. This kind of approach is vitally needed, especially in Lithuania (my country of origin). The implementation of River Basin Management Plans for the Water Framework Directive should start at the latest in 2012. Yet, until this day (the final stage of the preparation of River Basin Management Plans) there is no clue how the placements and types of different diffuse pollution abatement measures will be selected on a field level. There is only a vague understanding how much tons of nitrogen or phosphorus should be reduced per watershed. If no guidance is available, diffuse pollution abatement will result in random application of abatement measures, thus causing ineffective use of funds designated for the protection of water ecosystems. This work is designated for addressing this problem.

# 2 Goal, Objectives, Scope and Delimitations

## *2.1 Goal*

The main goal of this study is twofold. The first is to develop and apply a decision support method for the identification of critical source areas (CSA) for non-point sources and sensitive areas for pollutants in a selected watershed of the Lithuanian river. The second is to develop and apply a multi-objective spatial optimization (MOSO) technique for the selection and placements of best environmental practices (BEPs).

## *2.2 Objectives*

Three broad objectives have been raised in order to reach the goal of this study. First, to make an analysis of literature related to the identification of CSAs and the application of MOSO techniques for diffuse water pollution problem solution. Second, to prepare a selected physically-based, distributed or semi-distributed watershed model on the selected watershed in Lithuania. Third, to develop and apply methods for CSA, sensitivity identification and MOSO.

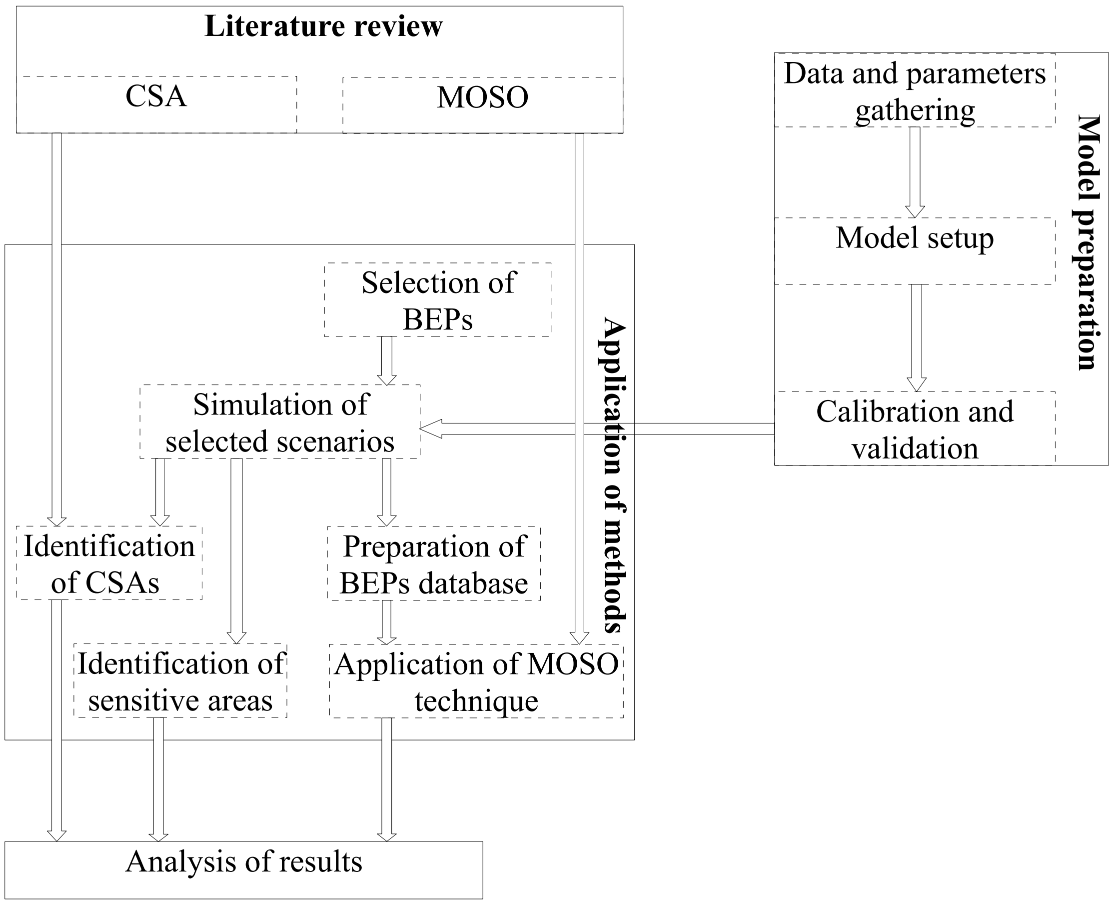
## *2.3 Scope/Delimitations*

This work is mainly intended to serve as a demonstration to watershed managers how analyzed problems could be solved. Model preparation and analysis parts were made according to literature and known practices. However, if the results should be applied for management purposes, more detailed local data are needed in regard to soil parameters, fertilization, tillage, economic values and wider range of BEPs should be considered as well. It is also important to mention that literature analysis on sensitivity identification techniques was skipped, because it was only a minor objective and ideas for it were quite clear from the beginning. Model selection and preparation have been build on the knowledge and skills gained from the previous work of Plunge (2009).

# 3 Work Flow

The sequence of activities done are presented in Picture 1. They are reflecting 3 main objectives raised for the study. Literature review and model preparation were done simultaneously. As model and area selection has been based on the previous work of Plunge (2009), it was possible to do the model preparation together with literature review. A literature analysis for CSA and MOSO themes (attached to diffuse surface water pollution problems) was made. The model preparation stage was the most time consuming. It was performed in steps. Firstly, the required datasets and parameters were gathered from various institutions, LEPA being the main source. Then, data were converted to the right formats, and missing parameters were estimated. This was needed for the model setup stage and running the model. Probably the most time consuming stage was model calibration. Although initial SWAT runs were showing satisfactory results and SWAT-CUP program helped a lot in calibration, many things had to be corrected before good results could be produced. The model was calibrated for water flows and nitrate loads. Model validation was done after calibration.

Literature review provided necessary knowledge and ideas for the application of methods stage. Specifically, literature review gave input to the identification of CSAs and the application of MOSO. The application of methods began with the selection of BEPs and the simulation of selected scenarios. Selection of BEPs was based on three criteria: suitability for simulation with the prepared model, possibility for installation on the selected watershed and availability of economic values. During simulation of scenarios, the selected BEPs as well as baseline and land response to nitrogen fertilization were evaluated . The results from simulation of scenarios provided input to the identification of CSAs, identification of sensitive areas and preparation of BEPs database. BEPs database consisted of loadings for each HRU under each scenario, which came from the simulation of scenarios with the prepared model. BEPs database also consisted of cost values for each HRU under each scenario. The application of MOSO technique was the last step in the application of methods stage. During this step a genetic algorithm was developed and used in the optimization of BEPs placement. The final results are shown in maps and the Pareto optimum graph. The analysis of results stage was used for the discussion of findings.



*Picture 1: Work flow of the study.*

# 4 Literature Review

Literature review was narrowed to the topics relevant to the study. Two questions were examined. One is the identification of CSAs for the purposes of diffuse water pollution management. Second is the application of MOSO for solving questions of BEPs selection and distribution questions.

## *4.1 Identification of Critical Source Areas*

Understanding of the inland surface water diffuse pollution phenomenon requires to have good knowledge about processes occurring in a watershed. Spatial dimension is essential for this understanding. In order to effectively tackle diffuse pollution it is necessary to locate the most important source areas of pollutants effecting water bodies. They are often called critical source areas (CSAs).

### 4.1.1 CSAs Definition

In general CSAs are defined as areas, which “contribute most pollutant load of the entire watershed and have a decisive impact on the receiving water quality” (Ou & Wang, 2008). However to define CSAs in practice is not as straightforward. There are many possible ways, which were found in literature.

The first problem is how to define “a decisive impact” on the water quality in a watershed. In other words, how to transform qualitative terms into to quantitative terms. Scientific literature provides a number of examples for solving this problem. For instance, White et al. (2009) define CSAs for sediment and phosphorus yield by ranking discrete units comprising the watershed according to their predicted contribution to the total load of sediment and phosphorus yield. 2.5% and 5% of total loads were chosen as benchmarks for CSA identification. Tripathi et al. (2003) defined CSAs for soil loss as sub-watersheds where predicted soil loss exceeded the “tolerable” level of 11.2 t/ha per year, which was based on previous studies in that area. CSAs identification for nutrients was based on threshold values for nitrate nitrogen and for dissolved phosphorus loading to water bodies, which were 10 mg/l and 0.5 mg/l. Those values were obtained from US EPA Quality Criteria for Water of 1976. Sivertun & Prange (2003) divide CSAs into risk and sub-risk areas. Risk areas are defined as areas, which obtain values of loadings higher than the mean value by two standard deviations and sub-risk areas as areas, which obtain values higher than the mean value by one standard deviation.

Another important aspect in the definition of CSAs is the spatial scale or spatial elements for which CSAs are defined. In general, the identification of CSAs is desirable on as detailed scale as possible. However, this might be very tricky. Therefore in some studies (Tripathi et al., 2003, Ouyang et al., 2008) critical sub-watersheds instead of critical areas in the watershed are identified. In other studies (White et al., 2009, Srinivasan et al., 2005) hydrological response units[3] (HRUs) are the basis for CSAs identification. In GIS based methods a raster cell is often used as the unit on which calculations and CSAs identification is obtained. Examples of this can be found in the articles of Sivertun & Prange (2003) and Ou & Wang (2008).

It is also necessary to consider the perspective when defining CSAs. According to Mass et al. (1985), CSAs can be defined from the land resource perspective and from the water quality perspective. Land resource perspective emphasizes the importance of those areas where soil erosion

3 Hydrological response units represent areas within sub-basin with unique combination of soil, land use and slope that is simulated as a single unit (White et al., 2009).

is higher than could be tolerated. Water quality perspective points to the areas where the best management practices (BMPs) could achieve the greatest improvement with the lowest cost.

Lastly, it is important to mention that ideal criteria for CSAs definition should encompass hydraulic transport of pollutants to a watercourse, magnitude of pollutant source, type of water resource and type of pollutant (Watershedss, 2010). In some cases, the identification of CSAs for certain pollutants is based on CSAs on other pollutants (as carrying vectors) or on CSAs defined for generating surface runoff. For instance, areas with increased soil loss would definitely have increased loadings of sediments and nutrients on the surrounding water bodies (Sivertun & Prange, 2003). According to Srinivasan et al. (2005) areas with increased surface runoff could be used to identify CSAs for phosphorus. Some authors (Qiu, 2009, Rao et al., 2009) suggested that CSAs for phosphorus and related pollutants could be approximated by variable sources areas (areas that actively generate runoff).

### 4.1.2 Importance of CSAs

The importance of CSAs for diffuse pollution abatement is well stressed by many authors (Qiu, 2009, Trevisan et al., 2010, Strauss et al., 2007, Diebel, et al., 2008, Ouyang et al., 2008, Tripathi et al., 2003, White et al., 2009, Noll & Magee, 2009). The contribution of pollutant loads is not uniformly dispersed in watersheds. Some areas have much greater influence on water bodies than others. For example, in the study made by White et al. (2009) 5% of that land area was responsible for 50% of sediment loads and 34% of phosphorus loads. Another example is the study made by Diebel et al. (2008) on Hefty Creek watershed in Wisconsin (USA), which concluded that 26% of phosphorus loss reduction could be achieved by targeting conservation measures to only 10% of fields. White et al. (2009) concluded that loads from CSAs were 3 to 10 times higher compared to average loads from agricultural fields. Therefore, it is quite obvious that in order to increase effectiveness of BEPs, abatement measures should be targeted on CSAs.

The significance of CSAs has been recognized by governmental institutions as well. For instance identification of CSAs is required for the projects of the US Rural Clean Water Program (Watershedss, 2010). The requirement for CSAs identification is primarily related to the requirement for cost-effectiveness of the selected abatement measures. According to Gitau et al. (2004) effectiveness of BMPs is significantly related to their placement in a watershed. Thus, it is possible to conclude that CSAs identification is one of the most important steps in inland surface water diffuse pollution abatement.

### 4.1.3 Simple Methods for CSAs Identification

Methods for CSAs identification can be divided into two broad categories, i.e. simplified and advanced methods. Simplified methods or models usually just show possible risk areas, whereas expert models or on-site exploration provide a more detailed analysis as well as quantities of pollutants coming from CSAs (Sivertun & Prange, 2003). Simple methods are fast, relatively cheap, requires little data and can be easily applied on large areas. They are often used in a screening stage.

An example of the simplest method for CSAs identification is the designation of all croplands within a quarter mile of water bodies to CSAs (Watershedss, 2010). Also CSAs can be identified on a sub-watershed level based on how much of the watershed is covered by agricultural lands (Minister of Environment of the Republic of Lithuania, 2005). Two most widely known (and used in practice) simple CSAs identification methods are phosphorus index (PI) and universal soil loss equation (USLE) model. PI was developed by US researchers. This method uses fertilization rate by phosphorus, phosphorus in soil, size of livestock, population density, transport factor, soil erosion, distance to stream, surface runoff and other factors (depending on version of PI) to come up with the index number, which would allow the indication of increased phosphorus load areas (Ou & Wang, 2008). By overlaying different factors in GIS analysis it is possible to identify risk areas or CSAs for phosphorus. The final PI value has no physical meaning, however, it allows comparison between the areas. USLE allows to estimate annual average soil loss (in tonnes per acre or hectare) that occurs due to sheet or rill erosion (OMAFRA, 2000). Calculation of soil loss is done by multiplication of rainfall/runoff, soils erodibility, slope length-gradient, crop/vegetation and management, and support practice factors. There are different adaptations of the USLE model. Sivertun & Prange (2003) developed the GIS USLE model, which is used for the identification of risk areas in much the same way as PI. Huang & Hong (2010) combined soil conservation service curve number (SCS-CN), USLE and nutrient losses equations into GIS based empirical model and applied it to identification of nitrogen and phosphorus CSAs.

Other examples of simple methods are erosion index and sediment yield index (Tripathi et al., 2003). A fuzzy modeling technique or simple overlays in GIS can be used for the identification of CSAs. Professional judgment by conservation managers can also be applied for qualitative evaluation of CSAs (White et al., 2009). There are many other methods as well.

Despite many benefits of using simple methods for the identification of CSAs (such as simplicity of analysis, low cost and time saving), there are many drawbacks, which should be addressed. The majority of simple methods are not capable of providing physically meaningful quantitative results because they do not represent actual physical processes occurring in an environment (Srinivasan et al. 2005). Index values allow comparison between the analyzed areas, however the extent of problems is beyond grasp of these methods. Moreover, there are many other restrictions in the use of simple methods. For example, according to Sivertun & Prange (2003), limitations of the USLE model are the following: factors for this model are only valid for areas that are similar to the areas, which were used in the development of the USLE model; the model could only analyze erosive slope parts and not accumulative; the model is applicable only to straight slopes, concave and convex slopes should be sub-divided to be included into this model. These and other problems determined that the attention is being focused on advanced methods for CSAs identification.

### 4.1.4 Advanced Methods for CSAs Identification

Advanced methods are used after screening stage, ideally for areas, which was identified as the areas of concern. Main distinction from simple methods are complexity of them (more than a single equation). Those methods encompass on-site exploration[4] and expert models. Expert models for CSAs identification are physically based models suitable for inland surface water diffuse pollution modeling. There are many examples of them, however to review all them is outside the scope of this study. Since this work was build on previous work of Plunge (2009), which selected SWAT model among other tools for diffuse pollution modeling, further discussion would focus on the SWAT model application for CSAs identification.

The SWAT model is one of the most widely used and robust physically based models for the assessment of diffuse pollution problems (Gassman et al., 2007, Srivinasan et al., 2005). Moreover, it has been integrated into GIS environment of ArcGIS (proprietary software) and Map Window (open source software). This is the key characteristic for the usefulness of such tool in CSAs identification. The successful application of SWAT for the purpose of CSAs identification has been demonstrated by many authors (Ghebremicheal et al., 2010, Srinivasan et al., 2005, White et al., 2009, Tripani et al., 2003, Ouyang et al., 2007). Most of them concluded that the SWAT model is a suitable tool for directing management efforts in abating diffuse pollution.

4 Methods for on-site exploration were not futher discussed, because there was no possibility to apply them.

However, a few problems have been raised in the studies examining the application of the SWAT model for the CSAs identification. The most important one is that the SWAT model does not simulate overland routing of pollutants and runoff (Srinivasan et al., 2005). This is due to the model concept, which treats loads within the sub-basin identically, disregarding their position of origination (White et al., 2009). This simplification increased the model's speed greatly, however, it reduced its capability to represent an environment in more detail. The SWAT model is good for predictions on a watershed scale, however, it was not designed for pollutant routing predictions at the detailed level. Nevertheless, according to White et al. (2009) correction of this shortcoming is planned in future versions of the SWAT model. Meanwhile this problem can be reduced by reducing size of sub-basins during watershed configuration stage. Other problems are connected with spatial calibration and validation of the model, data needs, etc. However, despite of these problems the application of the SWAT model for CSAs needs is highly recommended by the scientific community (Srinivasan et al., 2005, White et al., 2009, Tripani et al., 2003, Ouyang et al., 2007).

### 4.1.5 General Problems Connected to the Use of CSAs

There are also some general problems connected with CSAs identification and the concept itself, which should be discussed. Firstly, it must be mentioned that although many studies have identified CSAs, it is very hard to find studies, which would validate their methods with field experiments or measurements. According to White et al. (2009), "there is no quantitative assessment of program effectiveness if CSAs are actively targeted". Another problem is that very few simple methods provide physically meaningful quantitative results. Even if some methods can do this, uncertainty of the results obtained is very large. On the other hand, advanced methods usually have no official guidance how models should be used in the identification of CSAs. Since watershed modelers prepare models with different assumptions, different parameter sets, etc., there is no guarantee that the same (or at least similar) results will be obtained by different watershed modeling specialist even if they use the same model on the same areas and for the same time period. Thus comparison between studies is complicated. Also the importance of hydrological pathway is rarely well integrated into the methods used for CSAs identification (Watershedss, 2010). This is causing questions about the validity of results. There are other problems as well. Nevertheless, CSAs and their identification are the key component on any plan to abate inland surface water diffuse pollution.

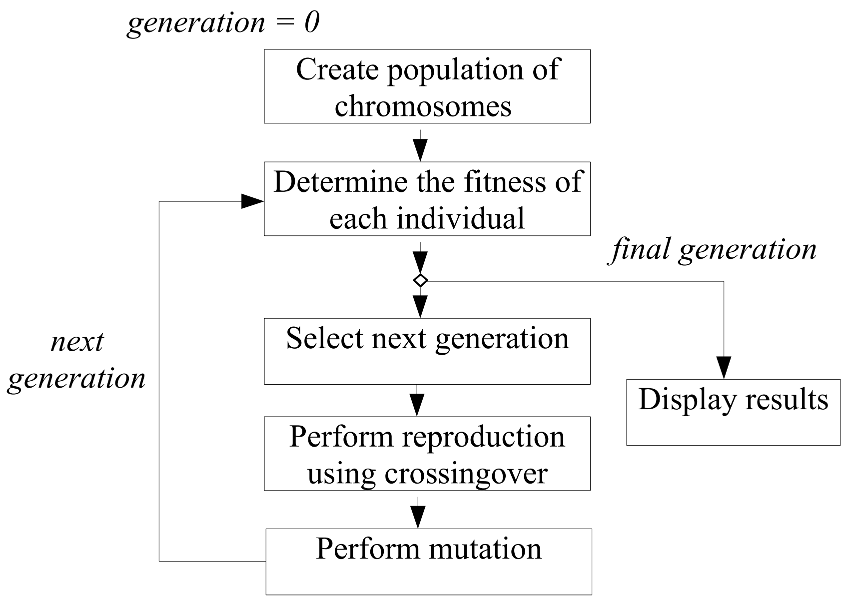
## *4.2 Multi-objective Spatial Optimization*

The effectiveness of BEPs for the control of diffuse pollution depends not only on the suitability of the site, but also on the right selection of BEPs. Those two factors affect the cost-effectiveness of diffuse pollution abatement programs the most. For finding an optimal distribution of BEPs as well as their optimal composition (in order to reduce abatement costs) multiple scenarios should be compared. However, this comparison is not possible with conventional methods. To find an optimal solution one should consider an exponential number of possible scenarios. According to Veith et al. (2003), for four nonmutually exclusive BEPs considered in 50 fields on some watershed, the number of possible scenarios would be $(2^4)^{50}$. The on-site selection of BEPs is neither practical nor economically feasible with such a number of possible options. In addition to this, field studies could not aid much in solving this problem. Establishing BEP effectiveness for a particular scenario takes many years. Yet, results are site specific and dependent on temporal variability in climatic conditions in that specific area (Veith et al., 2004). Moreover, multiple and often conflicting objectives should be combined when planning diffuse pollution abatement. Environmental, economical, sometimes institutional, esthetic and maybe even other types of objectives are equally important. Thus, finding solutions to such complicated problems seems to be an impossible task. No surprise that currently the placement of BEPs for abating agricultural diffuse pollution is done largely at random (Gitau et al., 2004, Maringanti et al. 2009). Yet, increasing numbers of studies have been done to solve this problem. They commonly apply techniques, which can be assigned to the group of multi-objective spatial optimization (MOSO) methods. Those methods integrate GIS, some kind of diffuse pollution modeling, an economical analysis and the most important, selected optimization methods.

Optimization problems are solved with many different methods: goal programing, linear programing, response surface methodology, shuffled complex evolution, simulation annealing, tabu search, genetic algorithms (GA) and others. These methods come from and are used in different disciplines: economics, engineering, informatics, mathematics, biology, evolution, etc. Despite this abundance of methods, most of the reviewed studies (Veith et al., 2003, Veith et al., 2004, Gitau et al., 2004, Maringanti et al., 2009, Arabi et al., 2006, Jha et al., 2009) used GA for the BEPs' optimization purposes in watersheds. For all of these studies GA usage was successful. The effectiveness and easiness of the MOSO problem formulation makes GA the most common choice. Therefore, only this method will be presented in more detail.

### 4.2.1 Genetic Algorithms

Genetic algorithms (GA) is the group of methods inspired by evolutionary biology. They are using the same principles of selection, inheritance, crossover and mutation, as life used for its evolution. Picture 2 presents the example of basic framework for GA, which can be also applied to watershed problems.




*Picture 2: General framework for GA (EDC, 2010).*

Each scenario for a watershed in GA is represented by a chromosome. A chromosome represents an individual in a population. A chromosome is composed from genes. Genes have information about the choice of management options for a certain field (it could also be a HRU or a sub-watershed). Each gene has possible allele sets. They are alternative management options for the same field. A selected number of initial individuals (represented by chromosomes) is randomly generated. During generations individuals with the highest fitness score are selected for breeding. During breeding the chromosomes of two individuals exchange parts through the process called crossover and two new chromosomes are formed. During each generation the lowest rated individuals (based on their fitness score) are removed from the population. In this way, the optimal solution is found after many iterations (generations in GA terminology). However, it is possible that this method would end up in local optima instead of global optima. Therefore, a mutation (random changes of genes) is introduced. Its probability should be low enough to keep the optimization results within the population, yet large enough to prevent convergence to local optima. The mutation rate for watershed studies of 0.01 has been proposed by Gitau et al. (2004) and Veith et al. (2004), whereas Maringanti et al. (2009) found the optimal gene mutation rate for its study to be at rate of 0.001. Other important parameters such as population and generations number greatly varies between mentioned studies, while optional parameters such crossover probability[5] and replacement rate[6] are based on the same values. All GA parameters of the mentioned studies are presented in Table 1. Even if GA parameters are specific for each study, this information could be helpful to get ideas about the choice of initial parameters when preparing GA.

*Table 1: Optimal GA parameter sets found by different studies.*

| Parameter | Maringanti et al., 2009 | Gitau et al., 2004 | Veith et al., 2004 |
|---|---|---|---|
| Population | 200 | 15 | 15 |
| Number of generations | 40000 | -- | ~1600 |
| Crossover probability | 0.9 | 0.9 | 0.9 |
| Replacement rate | -- | 0.7 | 0.7 |
| Mutation probability | 0.001 | 0.01 | 0.01 |

Other important factors to consider when preparing GA are termination criteria, allele sets and the type of GA linking with other components of optimization. Termination criteria can be defined either by the maximum number of iterations or by minimal improvement in the maximum fitness score. Both criteria can be used at the same time. It may be also important to define possible allele sets (management options) for each land use type or HRU. This is essential, since each land use type has only a certain group of applicable BEPs, which could be implemented or might be considered for implementation on this land use type.

A GA linkage with other components may be static or dynamic. A static linkage uses results of other components not linking them directly during the GA optimization. Dynamic linkage usually integrates the pollution simulation model and GA in the same simulation. The benefit of dynamic linkage is that pollutants routing can be accounted in the optimization. However, the dynamic linkage makes the optimization so much slower, that it can be used just on very small watersheds. In the study of Maringanti et al. (2009), static linkage was used and routing was not considered in the optimization. Authors argued that the inclusion of in-stream processes was not significantly changing optimization results. Therefore, routing could be excluded from the optimization with the benefit of increasing optimization speed.

Lastly, it is necessary to mention that the term GA can refer to the optimization method and to the optimization program. Scientists often write the code for GA programs by themselves or use the

5 Crossover occurs with defined probability during breeding.

6 Only certain part of population is replaced.

prepared one. For example, in the study of Perez-Pedini et al. (2005) the commercially distributed GA program Evolver© (http://www.palisade.com/evolver/) was used. Gitau et al. (2004) and Veith et al. (2003) used the freely available GALib package (http://lancet.mit.edu/ga/).

### 4.2.2 Other Factors to Consider in an Optimization

A few other factors, which have not been mentioned above, should be taken into account when preparing MOSO. Optimization is usually done by the integration of several tools. One of them is GA, another might be the watershed model and the third one might be the BMP tool[7]. In the studies of Maringanti et al. (2009) and Gitau et al. (2004), the BMP tool has been used as a component incorporated into optimization. According to Miringanti et al. (2009), “the BMP tool is a database that contains the quantitative information regarding the effectiveness of a BMP to reduce a particular pollutant from a given land use”. The BMP tool of Gitau et al. (2004) study contained 32 BEPs, which, according to the authors, could be divided into eight classes, such as animal waste systems, barnyard runoff management, conservation tillage, contour strip crop, crop rotation, vegetated filter strips, nutrient management plans, and riparian forest buffers. BMP tools are developed in two ways. Either the BMP tool is based on the results from BEPs monitoring studies as is in Gitau et al. (2004) or it is made by employing models as in the study of Miringanti et al. (2009). Miringanti et al. (2009) used the SWAT model for the simulation of the baseline (scenario without BEPs) and then simulated each BEPs separately on the selected group of HRUs. From this information authors calculated the effectiveness of each BEP on HRUs for which this BEP could be applied.

One more component necessary for MOSO is BEPs costs information. Some kind of database of annualized cost should be incorporated into optimization. The previous section (on environmental versus economical objectives) presented the discussion about what information should be included into this component.

Another important factor, which should be considered in MOSO, is how final results will be obtained. Is it by using a single objective function or Pareto optimum? The single objective function has an advantage of providing one answer that is straightforward to translate to spatial dimension. Whereas Pareto optimum provides multiple solutions, which are equally good. Thus, it is not as straight forward to provide one nice looking final map. However, according to Miringanti et al. (2009) these multiple near-optimum solutions are a great advantage over the single-objective function, since it allows decision makers to get insight into trade-offs between different solutions. The final result can be translated to map after decision makers decide, which solution is suitable for them.

Finally, it is important to mention the influence of uncertainties in MOSO. It is generally ignored by most of the studies. No surprise, since it would complicate the optimization process even further. However, according to Meyer et al. (2009), the influence of uncertainties is largely overlooked and the exclusion of them from the analysis could introduce far higher errors in the final results than one might expect. For more discussion on uncertainties and the use of uncertainty analysis in watershed studies please refer to Plunge (2009).

### 4.2.3 Relation to CSA

Before examining MOSO methods and their use, it is important to understand their relation with CSA, since it is another important part of this work. In the article of Veith et al. (2004), this link has

7 It should be called the BEP tool in the European context. However articles, which mention it, are only American.

been provided. Authors divide all methods for the selection of BEPs applications into plan-based and performance based methods. Plan-based methods are built on past field studies and scientific theory. This category includes targeting methods, which direct diffuse pollution abatement towards CSAs. On the other end there are performance based methods, which apply simulation models to assess the effectiveness of various scenarios of BEPs application. Optimization methods are assigned to the latter category. Those methods are applied with different approaches for diffuse pollution reduction. For instance, an incremental approach is based on the idea of identifying CSAs and then targeting the resources towards the most important areas, adding less critical areas in the future when more funds are available (Perez-Pedini et al., 2005). This systematization of methods is quite straightforward, yet sometimes can be mixed up. For instance, models are used for CSA identification as well. Nevertheless, this system of classification is useful in understanding the relationship between CSAs and MOSO.

Targeting methods are probably most popular between watershed managers. Yet, according to many authors, optimization methods provide more cost-effective solutions. For instance Arabi at al. (2006) found that optimization reduced the cost of diffuse pollution abatement two times compared to the targeting plan. Veith et al. (2004) found that optimization also achieved better cost results than targeting. On average costs were reduced by 15%. According to Veith et al. (2004) optimization also includes spatial interaction between BEPs (which is not possible for targeting). It also offers more flexibility in the choice of placement and selection of BEPs to achieve the required reduction of pollution. Yet, targeting has its benefits too. Its results are simpler to implement and interpret, and it requires less information compared to optimization. Therefore, the trade-off between the potential benefits and drawbacks should be weighed before selecting any method for the analysis.

### 4.2.4 Environmental versus Economic Objectives

One of the most important advantages of MOSO methods is their ability to combine environmental and economic (and other) objectives in one result. These methods provide stakeholders and decision makers with solutions, which take into regard much of their concerns. Moreover, MOSO methods can provide multiple near-optimum solutions, from which stakeholders may choose the most appropriate one for their case. A method called trade-off frontier or Pareto optimum is used for this purpose (Jha et al., 2009). This method puts all near-optimum solutions on one graph, where one axis usually represents the cost of solution and the other - potential reduction of pollution.

Pollution reduction potential is usually calculated with some kind of model. For instance, Meyer et al. (2009) used SWAT to calculate nitrogen leaching potential, which was applied in the optimization. Veith et al. (2003) used USLE to calculate sediment loads used in the optimization. If the optimization is concerned with more than one pollutant, weights are used (according to stakeholder priorities) to produce some kind of pollution reduction potential index (Jha et al., 2009). Optimization is often done by achieving environmental objectives first, and then trying to reduce the costs of pollution abatement. It is so called the two-part fitness equation (Gitau et al., 2004). The maximum acceptable level of diffuse pollution is set to the stakeholder agreed value (Veith et al., 2004). Then each solution, which passes that threshold, is passed to the calculation of costs. The least costly solution is selected in the end. Yet, in some cases another way could be preferred. Target costs (available funds for some pollution abatement program) could be used as a basis for identifying what solution could be passed to fitness evaluation (which in this case would be based on pollution reduction potential).

The calculation of pollutant reduction scenarios requires much skills and effort. Yet, cost calculation is quite complicated as well. The total cost of BEPs should include implementation and maintenance costs just as basic information. Also incorporation of opportunity costs, which refer to the costs of not choosing the management practice with the highest return, is quite important (Veith et al. 2003). So is the use of the discount rate for the calculation of future costs and benefits of BEPs (Jha et al., 2009). Additional costs mentioned by Veith et al. (2003) are the following: the public cost of contracting (costs incurred while creating agreements with farmers for the change of their management practices), enforcement (costs to ensure that agreements with farmers are met), reimbursement costs (if farmers are compensated for certain management actions) and information costs (costs used to generate optimal or near-optimal solutions). Moreover, since the lifetime of BEPs and their maintenance costs vary between BEPs and with time, it is important that the final costs of BEPs would be annualized (Gitau et al., 2004). Taking into account the mentioned complexity of cost calculation, it is important to state that a skilled economist is no less important for MOSO than a skilled watershed modeler.

However, dealing with economical and environmental objectives is not nearly enough. According to Veith et al. (2003), solutions should “conform to reasonable farming practices”. If the whole watershed is considered for an optimization, more requirements could arise and thus should be addressed in the formulation of the objective function.

## *4.3 Summary of Literature Part*

Literature review and analysis could be summed up with a few conclusions. Firstly, it is important to state that the identification of CSAs and the application of MOSO are both very important ways of dealing with diffuse pollution. Both are designed to aid decision makers. Each of them has certain benefits and drawbacks, and each of them has their appropriate application areas. The identification of CSAs a is simpler way, which allows to locate areas responsible for the deterioration of a water body status. Questions such as how to define CSAs, are qualitative results required, what scale is required for the results, etc., would shape CSAs identification and its complexity. The identification of CSAs requires less data; it is also easier to implement, interpret and use its results. However, CSAs cannot answer a very important question to decision makers dealing with watershed management. That is what and where should be done to reach the objectives with the least cost. CSAs can only be useful as a guide for identifying areas, which need certain attention. On the other hand, the application of MOSO is capable of answering the mentioned question. Yet, data requirements, qualification of specialists (at least on the watershed modeler and the economist), time needed to build MOSO, requirements for computational resources are far higher than in CSAs identification. The presentation and interpretation of the results can also be more complicated. Furthermore, the biggest difference in the level of uncertainty, which for MOSO application is considerably higher. Nevertheless, no simpler alternative is capable of providing answers to the questions, which are key to diffuse pollution abatement. There are many methods for solving MOSO problems, yet GAs were the only choice in the reviewed studies dealing with diffuse water pollution problems. The application of MOSO can be designed to provide different results. It can give a single best solution or many near optimum solutions. For the single best one, studies are using two-part fitness equations. For many near optimum solutions, methods providing Pareto optimum are used. It is also important to mention that in MOSO application GA's and the watershed model requires integration. This can be done through static (routing excluded) or dynamic (routing included) linkage. Static linkage is much simpler and requires less computational resources. In MOSO application for diffuse pollution problems static linkage has more advantages compare to dynamic linkage. Lastly, it is necessary to emphasize the importance of correct information on the costs of diffuse pollution abatement (public and private) and on the benefits in MOSO application.

# 5 Method Application

This section presents the preparation of the model, the methods for CSAs and sensitive areas identification and MOSO application.

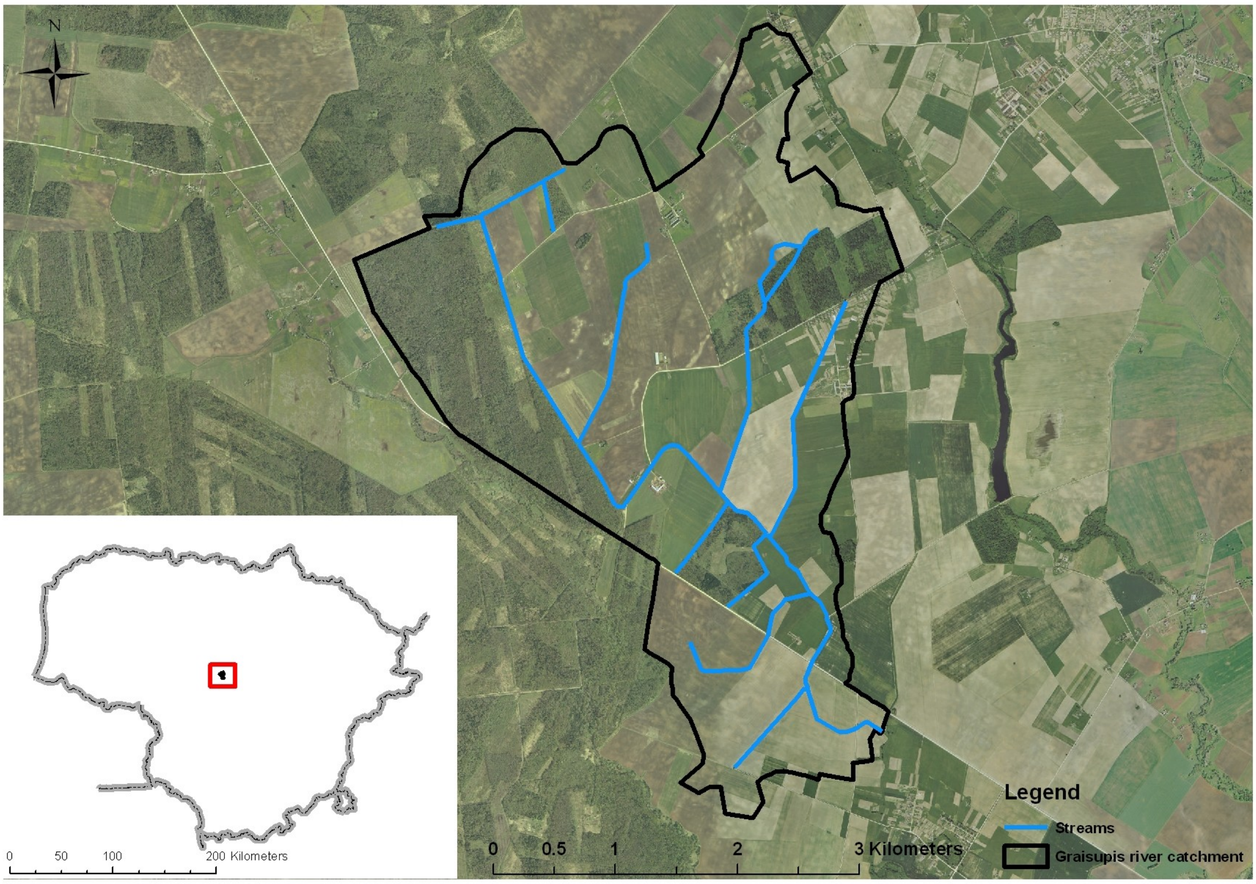
## *5.1 Model Selection*

Model selection is based on the previous work done by the author of this thesis. In the Master thesis of "Risks versus Costs: A New Approach for Assessment of Diffuse Water Pollution Abatement" (Plunge, 2009) watershed models have been examined based on many criteria. The SWAT model was chosen as the most suitable watershed model due to many reasons. To mention a few of them: suitability for diffuse pollution assessment, physically based and semi-distributed parameters, the possibility of the evaluation of long time periods, reliability, integration with GIS, etc. Moreover this model has a huge community of scientists and professionals working with it and numerous application examples published in scientific literature. This makes this model one of the most sound tools to use in any study related to the assessment of difuse water pollution problems. The requirements for modeling a tool for this study were similar as for the previous one. Therefore, the SWAT model has been chosen. The interested reader is refered to SWAT theoretical documentation (Neitsch et al., 2005) for details on the model. The exact version of the SWAT model used in the study was ArcSWAT 2.1.6.

## *5.2 Area Discription*

The area selected for modeling was the Graisupis river catchment. It is located in the center part of the Republic of Lithuania (see Picture 3). The total area of the Graisupis river catchment is only around 14.2 square kilometers. This area lies in the Lithuanian Middle Plain, which is dominated by fertile soils. Agricultural areas and pastures dominate this part of Lithuania. This is common for the Graisupis river catchment as well. 71% of the Graisupis river catchment is occupied by agricultural areas and pastures. Forests occupy the rest 28% with 1% leaving for build-up areas and water bodies. The catchment is situated 57-70 meters above sea level. It receives around 608 mm of precipitation annualy. The Graisupis river catchment is dominated by Gleyic Cambisols soil type. There is a small settlement of Azuolaiciai with 26 homesteads. There is also one larger cattle farm with around 200 animals. Most crops cultivated in the area are wheat, barley, maize, sugar beet and winter crops.

The main benefits from the selection of the Graisupis river catchment are several. Firstly, this catchment is unique in Lithuania for the length of different detailed monitoring activities performed on small agricultural areas. The monitoring activities of agro-ecosystems stretch at least to the year 1998. The Water Management Institute (WMI) of Lithuanian University of Agriculture is responsible for these monitoring activities. The monitoring activities are payed by and the data are supplied to the Lithuanian Environmental Protection Agency (LEPA). Secondly, the Graisupis river catchment is dominated by agricultural areas, which are the main contributors to diffuse water pollution. The agricultural activities in Lithuania is the major water deterioration factor (LEPA, 2010). Thirdly, the size of the Graisupis river catchment makes it very suitable as a test area for the application of new assessment methods. There are other minor benefits as well.




*Picture 3: Location of the Graisupis river catchment (all the maps presented in this work apply Lithuanian Coordinate System (LKS-94)).*

## 5.3 Data and Parameters

Most of the data required for the study were obtained from the LEPA. The LEPA organizes many monitoring activities itself, yet even more data are obtained through exchange agreements with other governmental institutions or procured from commercial bodies. Only meteorological data were provided by the Lithuanian Hydrometeorological Service (LHMS) under the Ministry of Environment. Soil parameters were obtained from the global soil database.

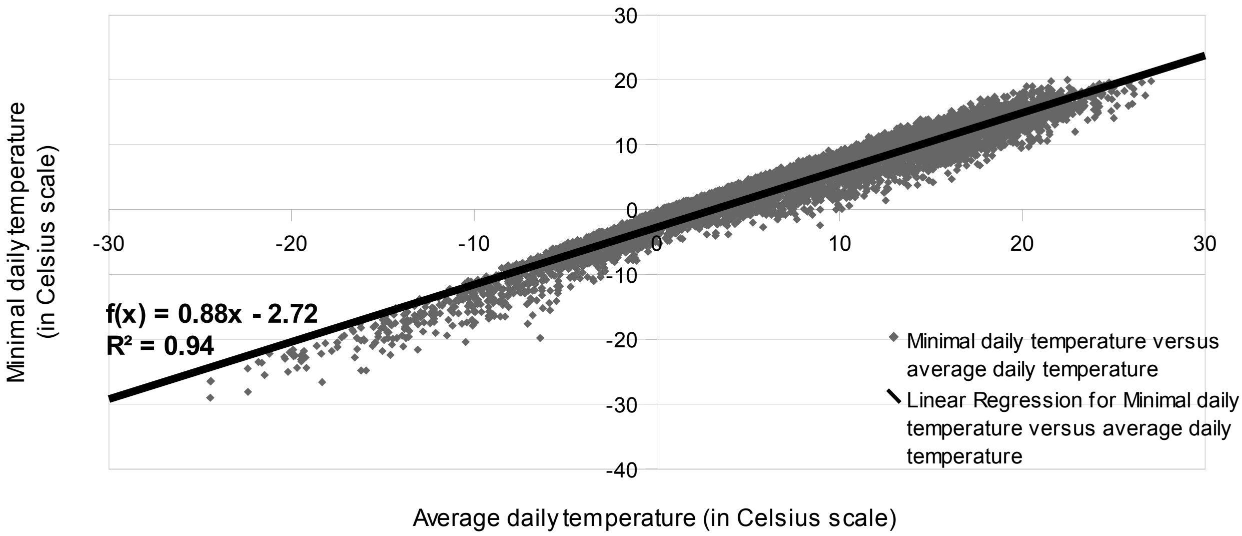
### 5.3.1 Data Inputs

The point elevation model (PEM) was used in the preparation of the digital elevation model (DEM). The resolution of the PEM was 5 meters. It was the most recent and detailed elevation data at the time of the study. The elevation data were recorded in year 2005 and were managed by the National Land Service under the Ministry of Agriculture of the Republic of Lithuania. Thus, this database was chosen as the most suitable one. The PEM was stored in files, which could not be read by ArcGIS 9.2 software. Therefore a small script in Matlab was written to convert PEM files to text files importable into ArcGIS. The prepared DEM in presented in a part of Picture 6.

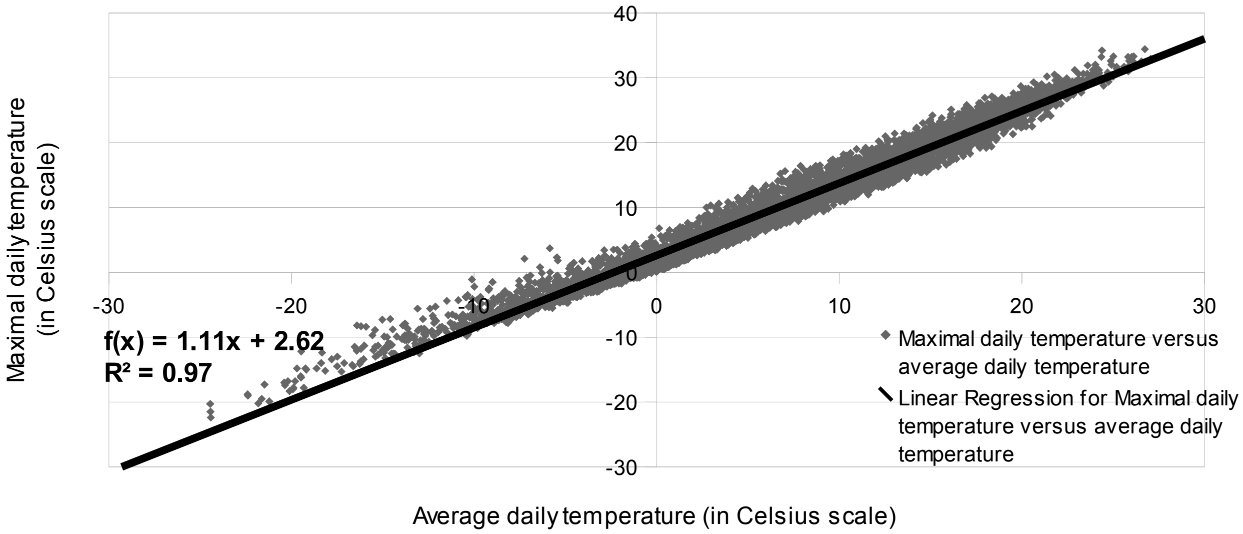
The monitoring data of land use, catchment borders, location of reaches, flow and water monitoring data were obtained from the WMI's reports "The Analysis of Land Use, Chemical Composition of Water and Precipitation in Typical Agro-ecosystems of Middle and West

Lithuania”. These reports are issued annually since 1997 for the LEPA. They supply the monitoring results of typical agricultural catchments for the purpose of gaining needed knowledge for the calculation of nutrient loads from agricultural lands. Monthly values of water hydro-chemical parameters for the period of 1998-2009 have been used to prepare an observation file for the SWAT model calibration and validation. Daily discharge from the catchment data for the period of 2000-2009 were used for the calibration and validation of hydrology. The meteorological data were obtained from the LHMS. 21 years of time series were requested from the LHMS for hourly temperature, precipitation, relative humidity, temperature dew point, solar radiation and wind speed. Most of the data were collected in the Dotnuva meteorological station (located 4 km to the North from the Graisupis river catchment), except for the solar radiation data, which were collected in the Kaunas meteorological station (located 47 km to the South from the Graisupis river catchment). There were no solar radiation measurements in the Dotnuva meteorological station. Only daily averages for temperature and wind speed have been obtained from year 1988 to 1993. Hourly data (every three hours) were obtained for the period of years of from 1993 to 2008. Daily average of relative humidity and temperature dew point data were obtained from 02/1992. Hourly data (for every three hours) for relative humidity was available from year 1993 and for temperature dew point from 12/1993. Precipitation was available as daily cumulative samples from 1988 to 2007, and as hourly (for every six hours) samples from 2007. Solar radiation data were obtained from year 1998 to 2008. Daily minimal and maximal temperature, precipitation, wind speed, relative humidity and solar radiation were used as input of weather data time series for the SWAT model.

All inputs to the model for the weather time series should be of the same length. The time period chosen for the model input was 1988-2008. Solar radiation and relative humidity were not as long. Therefore, the value of -99 was used for the input from year 1988 until the measurements started. The value of -99 calls the weather generator module within the SWAT model. The weather generator uses statistical data obtained from weather measurements to generate the missing weather data. Temperature data for the SWAT until 1993 were estimated from the relationship obtained from hourly temperature data for the period of 1993-2008 (see Picture 4 and Picture 5). As the determination coefficients for both equations were high enough (more than 0.94), the estimation method was assumed good enough to provide the needed data.




*Picture 4: Relationship between minimal daily temperature and average daily temperature.*




*Picture 5: Relationship between maximal daily temperature and average daily temperature.*

Statistical data for the weather generator were prepared using the same weather monitoring data described in the section above. Average or mean daily maximum air temperature for month (TMPMX) expressed in degrees of Celsius, average or mean daily minimum air temperature for month (TMPMN) expressed in degrees of Celsius, standard deviation for daily maximum air temperature in month (TMPSTDMX) expressed in degrees of Celsius and standard deviation for daily minimum air temperature in month (TMPSTDMN) expressed in degrees of Celsius were calculated using the maximum and the minimum daily temperature data[8]. Average or mean total monthly precipitation (PCPMM) expressed in mm $H_2O$, standard deviation for daily precipitation in month (PCPSTD) expressed in mm $H_2O$ per day, skew coefficient for daily precipitation in month (PCPSKW), probability of a wet day following a dry day in the month (PR_W1), probability of wet day following a wet day in the month (PR_W2) and average number of days of precipitation in month (PCPD) were calculated from the precipitation data time series by using the pcpSWAT program, which was prepared by Stefan Liersch (http://www.brc.tamus.edu/swat/pcpSTAT.zip). Maximum 0.5 hour rainfall in entire period of record for month (RAINHHMX) expressed in mm $H_2O$ was the only one parameter, which was impossible to obtain or calculate from the available meteorological data. An assumption was made that the maximum 0.5 hour rainfall for the entire period of record for each month is more or less equal to the six hour cumulative rainfall sample for the period of 2007-2008. Average daily solar radiation for month (SOLARAV) expressed in MJ/m$^2$/day, average daily dew point temperature in month (DEWPT) expressed in degrees of Celsius and average daily wind speed in month (WNDAV) expressed in meters per second were calculated from daily average values obtained from the LHMS. All statistical parameters are presented in Table 2.

8 Part of data 1988-1993 were estimated from statistical relationship with average data.

*Table 2: Statistical parameters for the weather generator.*

| Month | TMPMX | TMPMN | TMPSTDMX | TMPSTDMN | PCPMM | PCPSTD | PCPSKW | PR_W1 | PR_W2 | PCPD | RAINHHMX | SOLARAV | DEWPT | WNDAV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | -0.03 | -4.22 | 5.42 | 6.14 | 35.49 | 2.13 | 3.72 | 0.29 | 0.64 | 14.81 | 12.7 | 1.79 | -4.87 | 2.96 |
| 2 | 0.22 | -4.81 | 5.28 | 6.31 | 31.49 | 1.84 | 2.42 | 0.38 | 0.61 | 14.71 | 7 | 4.09 | -5.60 | 2.82 |
| 3 | 4.11 | -2.08 | 4.49 | 4.37 | 31.62 | 2.09 | 3.45 | 0.28 | 0.56 | 13.00 | 6.4 | 8.87 | -3.29 | 2.69 |
| 4 | 11.43 | 3.09 | 5.58 | 3.80 | 29.70 | 2.33 | 3.74 | 0.21 | 0.55 | 10.05 | 5.2 | 13.31 | 1.55 | 2.58 |
| 5 | 17.24 | 7.54 | 4.67 | 3.52 | 42.58 | 3.39 | 4.81 | 0.23 | 0.56 | 11.00 | 16.7 | 18.14 | 5.84 | 2.31 |
| 6 | 20.42 | 11.12 | 3.80 | 2.80 | 53.99 | 3.93 | 3.54 | 0.34 | 0.53 | 12.86 | 19 | 17.86 | 10.01 | 2.03 |
| 7 | 23.00 | 13.40 | 3.95 | 2.57 | 67.88 | 5.17 | 3.69 | 0.30 | 0.54 | 12.86 | 15.3 | 18.07 | 12.70 | 1.98 |
| 8 | 22.14 | 12.59 | 3.73 | 2.62 | 63.95 | 4.52 | 3.90 | 0.29 | 0.62 | 13.90 | 25.3 | 14.29 | 12.30 | 1.91 |
| 9 | 16.56 | 8.39 | 3.81 | 3.26 | 43.69 | 3.53 | 4.37 | 0.26 | 0.56 | 11.67 | 13.7 | 10.47 | 8.49 | 2.12 |
| 10 | 10.20 | 4.13 | 4.22 | 3.88 | 52.31 | 3.66 | 3.75 | 0.29 | 0.57 | 13.33 | 10.2 | 4.83 | 4.73 | 2.26 |
| 11 | 3.54 | -0.63 | 4.60 | 4.74 | 39.20 | 2.74 | 3.63 | 0.33 | 0.54 | 13.24 | 6.8 | 1.96 | 0.05 | 2.52 |
| 12 | 0.11 | -3.86 | 5.03 | 5.79 | 40.71 | 2.39 | 2.78 | 0.36 | 0.60 | 15.14 | 11 | 1.18 | -3.70 | 2.58 |

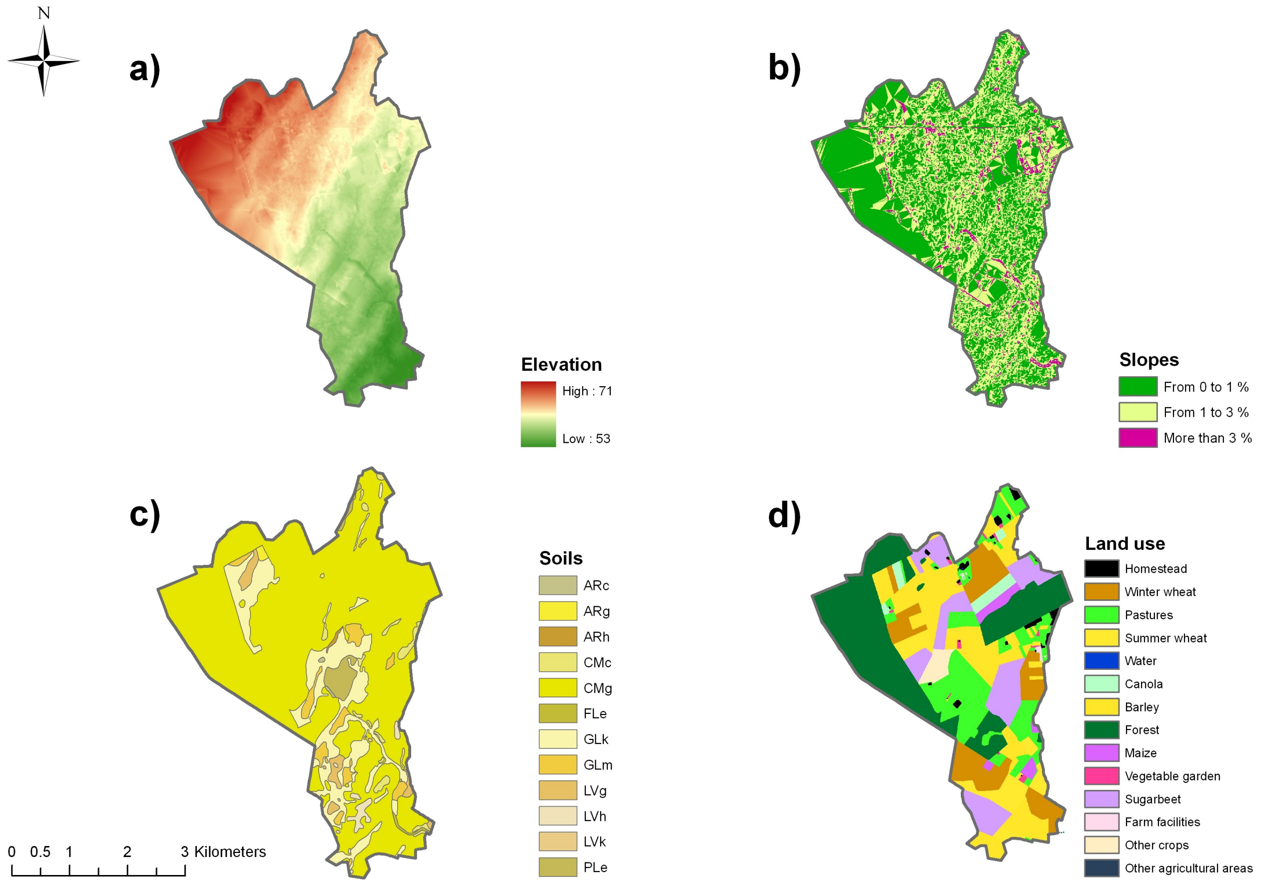
Two soil databases were obtained from the LEPA to prepare input for the SWAT model. The main database used for input was prepared by the National Land Service under the Ministry of Agriculture of the Republic of Lithuania. The scale of this GIS vector database is 1:10000. Another soil database of scale 1:300000 was used in the places where a more detailed GIS database lacked cover (mostly in areas covered by forests). The national soil classification system was linked to the Food and Agriculture Organization (FAO) soil classification. By linking it to the FAO system it was possible to get soil parameters from soil databases, which were prepared for the world. The Harmonized World Soil Database (HWSD) (Fischer et al., 2008) was used to obtain most of the soil parameters. Other parameters were estimated, left default or obtained through calibration. The estimation of soil parameters is explained in the next section. The prepared soil layer is presented in c part of Picture 6.

Land use in the Graisupis river catchment is monitored yearly by the WMI. Land use of 2008 was used as input for the model (see d part of Picture 6). This was due to the availability of data at the beginning of the model preparation. The most appropriate parameter compositions were assigned to land use categories from the prepared crop and urban SWAT databases through a look-up table. The assigned categories are presented below in Table 3.

*Table 3: Assigned land use from SWAT database comparing to original land use.*

| Original land use | Assigned land use from crop or urban databases | Code of assigned land use |
|---|---|---|
| Homestead | Residential-Low Density | URLD |
| Winter wheat | Winter Wheat | WWHT |
| Pastures | Pastures | PAST |
| Summer wheat | Corn | CORN |
| Water | Water | WATR |
| Canola | Spring Canola-Polish | CANP |
| Barley | Spring Barley | BARL |
| Forest | Forest-Deciduous | FRSD |
| Maize | Sweet corn | SCRN |
| Vegetable garden | Garden or Canning Peas | PEAS |
| Sugarbeet | Sugarbeet | SGBT |
| Farm facilities | Industrial | UIDU |
| Other crops | Agricultural Land-Close-grown | AGRC |
| Other agricultural areas | Agricultural Land-Generic | AGRL |

Slopes were obtained by using slope definition module in ArcSWAT. 3 categories have been chosen: from 0 to 1 %, from 1 to 3 % and slopes more than 3 % slopes. The distribution of slopes is presented in part b of Picture 6.




*Picture 6: Elevation (a), slope (b), soil (c) and land use (d) data for the Graisupis river catchment.*

## 5.3.2 Estimation of Soil Parameters

The obtained GIS soil database had very few (in most cases none) usable parameters for the model input. Usable data were the soil classification according to the national soil classification system, which was related to the FOA soil classification system. Overall 12 soil types were present in the Graisupis river catchment (Appendix, Table 1). These data were obtained from the soil GIS layer attributes. Soil parameters were mainly obtained from HWSD. Some parameters were estimated or left default. The SWAT model requires defining some parameters for the whole soil profile as well as for each layer in the soil profile. In the HWSD, the soil profile is divided into two soil layers: 0 to 30 cm and 30 cm to 100 cm. Therefore, the number of layers (LAYERS) for all soils was set to 2. The soil hydrological group (HYDGRP) was estimated using the USDA proposed classification, which is based on soil textures (such as sand, loam sand, heavy clay, etc, which were available in the HWSD). The maximum rooting depth of the soil profile (SOL_ZMX) was set to 1000 mm (as the overall depth of the soil profile) with the exception of Calcaric Arenosols, which was set to 700 mm because in the HWSD the obstacles to roots were set to 60-80 cm. The parameters for the fraction of porosity (void space) from which anions are excluded (ANION_EXCL) and the potential or the maximum crack volume of the soil profile expressed as a fraction of the soil volume (SOL_CRK) were left default at 0.5. These parameters are optional and the model proposed value (if no data is entered) is 0.5. The depth from the soil surface to the bottom of layer (SOL_Z), moist bulk density (SOL_BD), available water capacity (SOL_AWC), organic carbon content (SOL_CBN), clay content (CLAY), silt content (SILT), sand content (SAND), rock content (ROCK) and electrical conductivity (SOL_EC) were taken from the HWSD. Available water capacity was available just for all soil profile. Thus, an assumption was made that it is the same for both soil layers. Saturated hydraulic conductivity (SOL_K) (mm/hr) was calculated using Equation 1, which was developed by assessing the relationship between parameters in the prepared SWAT soil database. The strength of statistical relationship was $R^2$=0.76 between the values estimated by this equation and the real values.

$$K_s = \frac{1.7499 \times \mathrm{e}^{0.0751 \times SAND} + 1000 \times e^{-0.211 \times CLAY}}{2}$$

*Equation 1: Formula for calculation of saturated hydraulic conductivity.*

Moist soil albedo (SOL_ALB) was not available in the HWSD. It was estimated by using the colors of layers in the soil profile. Color codes in the Munsell color system for different types of soils were available in ISRIC-WISE Harmonized Global Soil Profile Dataset (Batjes, 2008). Soil profiles of the soil types existing in the Graisupis river catchment (and which are in Lithuania or closest to Lithuania) were selected for obtaining the colors of soil layers. Albedo was estimated by using the relationship between the (moist) soil color of the topsoil and the soil albedo given in the article of Gisjman et al. (2007). Color codes were related to the names of colors by using visual senses with the aid of website http://www.it.lut.fi/research/color/demonstration/demontration.html. The final result is presented in Table 4.

*Table 4: Albedo values assigned to Munsell color system codes.*

| Munsell color system codes of selected profiles taken from WISE database | Soil color | Albedo |
|---|---|---|
| 10YR3/4, 10YR3/3, 10YR4/3, 10YR2/1 | Black | 0.09 |
| 10YR5/4, 5YR4/4, 2.5Y5/6, 2.5Y6/8, 2.5Y6/4 | Brown | 0.13 |
| 10YR6/1,10YR5/2 | Grey | 0.13 |
| ----- | Red | 0.14 |
| ----- | Yellow | 0.17 |

The USLE equation soil erodibility factor (USLE_K) was estimated using the Williams equation given in the SWAT Input/Output File Documentation (Neitsch et al. 2004). This equation requires sand, silt, clay and organic content parameters. All soil parameters used in the model are presented in Appendix A.

### 5.3.3 Data Adjustments

After studying the SWAT technical documentation some problems were identified with the input data. The first was that the precipitation data might be incorrect due to the effects of rain gauges, which were not designed to shield wind effects. The design of rain gauges has been confirmed by Natalija Gaurilcikiene, the LHMS responsible person for the Dotvuna meteorologic station. It has been confirmed that the Dotvuna meteorological station uses only simple rain gauges, which are not designed to shield wind effects. According to Larson and Peck cited in the SWAT documentation for these types of rain gauges the deficiencies of 10% for rain and 30% for snow are common (SWAT technical documentation). Therefore, all precipitation data were changed based on this information. For days when there average temperatures were above 0 degrees of Celsius, the precipitation was increased by 10%, and for days when temperature was equal to or below 0 degrees of Celsius, the precipitation was increased by 30%.

Another problem with the input data was connected with wind speed measurements. According to the SWAT technical documentation, "SWAT assumes wind speed information is collected from gauges positioned 1.7 meters above the ground surface" (Neitsch, et al., 2005). However, wind speed measurements are done at 10 meters above the ground in the Dotnuva meteorological station. Therefore, a transformation was done with Equation 2 to obtain input required by the SWAT model. This equation has been approved and recommended by the LHMS (Smitiene, 2007).

$$V_{1.7\mathrm{m}} = V_{10\mathrm{m}} \times \sqrt[4]{\frac{1.7}{10}} \approx 0.642 \times V_{10\mathrm{m}}$$

*Equation 2: Formula for recalculation of wind data (wind shear exponent is equal to 0.25).*

## 5.4 Model Preparation

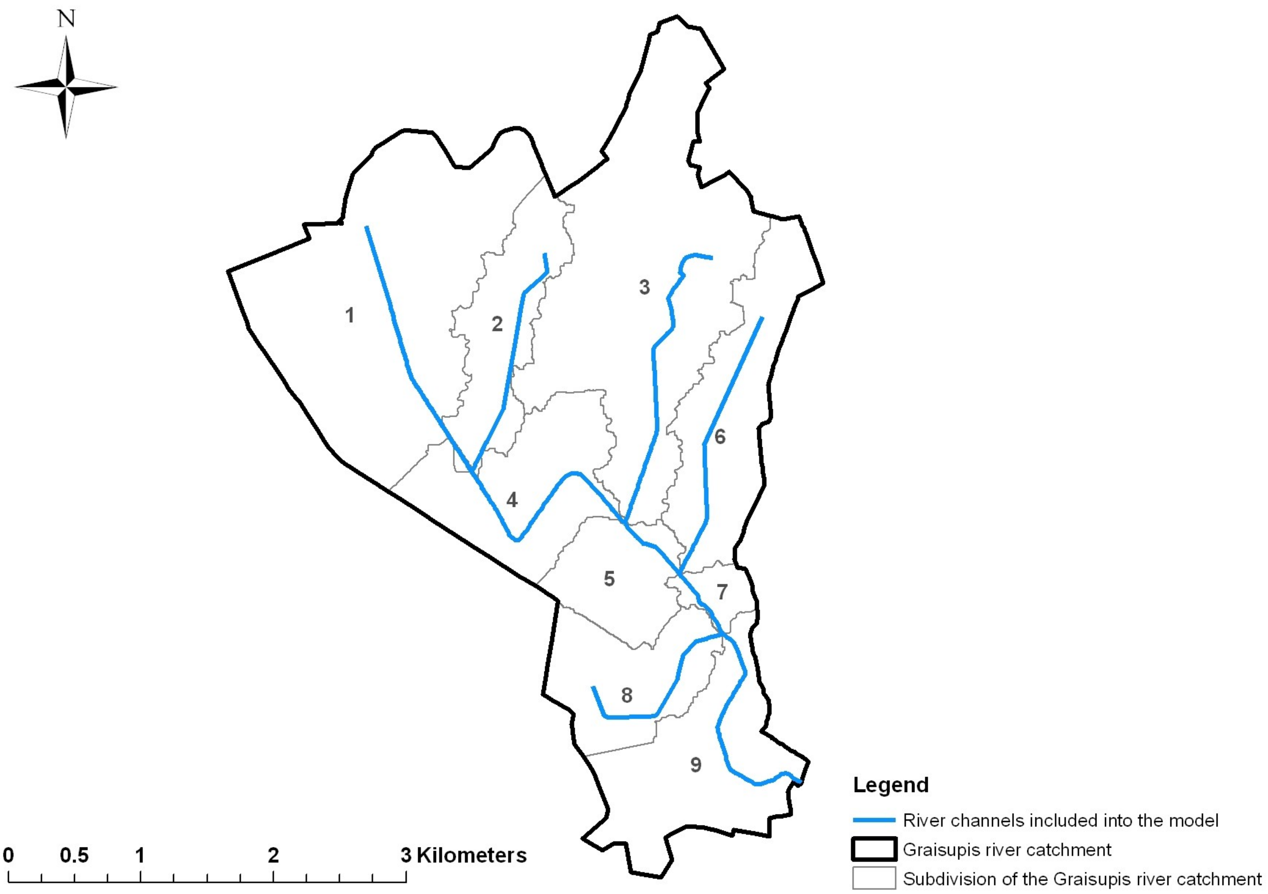
Model preparation was made by putting in the data and parameters into the right formats and databases, setting up the SWAT model with the use of SWAT modules and then calibrating and validating the prepared model. Calibration was the most prolonged faze. It was made with the help of the SWAT-CUP 2.1.5 program.

### 5.4.1 Dividing the Catchment

The SWAT model is a semi-distributed parameters model. Pollutant loads originating anywhere in a sub-basin are treated equally. Thus, it is better to divide the basin into smaller sub-basins to represent the spatial variability of physical conditions in a watershed. Moreover, shortcomings of the SWAT model for overland pollutant routing are reduced if more or all physically meaningful sub-basins are used.

For this study the Graisupis river catchment was divided into 9 sub-basins. For each of this sub-basin separate reach segment was assigned. The division of the catchment is presented in Picture 7.

For the watershed delineation the real boundaries of the catchment was used, which were obtained from the WMI report of 2009. The location of reaches was also obtained from the same report. The boundaries of the catchment were used as a mask and the reaches were burned into DEM with the watershed delineation module of ArcSWAT software. For DEM based stream definition an 80 ha area was chosen. With the mentioned settings the watershed delineation module of ArcSWAT software divided the whole catchment into 9 sub-basins. To align the external boundaries of these sub-basins with the real catchment boundary, the sub-basin layer was edited manually. Stream definition was made with the pre-defined watershed and stream datasets in order to set up the final watershed configuration.




*Picture 7: Division of the Graisupis river catchment.*

## 5.4.2 Calibration of Flow

For the calibration of water flow, daily data for the Graisupis river flow were available from year 2000. Monthly data for river flow were available from year 1998. The period from 2000 to 2006 was used for calibration. For validation purposes the data from year 2007-2009 were used. The split between calibration (7 years for calibration) and validation (3 years for validation) years was based on the common practices observed in scientific literature. For nitrate calibration monthly data from 1998 was available. Yet, calibration and validation of nitrate loads were made on the same periods as that water flow. The data of 1998-1999 were left out due to the lack of daily water flow values. The watershed model performance criteria were based on the article of Moriasi et al. (2007) (see Table 5 on the next page). The least aim for calibration was to reach a good performance rating. It was intended to use only monthly values in later steps. Thus, the statistics for a monthly time step were used.

The warm-up (or spin-up period as it is used in some literature sources) period of 3 years was selected for hydrology calibration and validation. For nitrates this period was 5 years. The length of warm-up periods was based on the time necessary to eliminate the effects of initial model values on final results. Several runs were performed with different warm-up periods. Those, which indicated the best effect on model performance increase, were selected.

*Table 5: General performance rating values recommended by the article of Moriasi et al. (2007) for statistics for a monthly time step.*

| Performance Rating | Root mean square error divided by standard deviation of measured data (RSR) | Nash-Sutcliffe efficiency[9] (NSE) | Percent of bias (PBIAS) | | |
|---|---|---|---|---|---|
| | | | Streamflow | Sediment | N, P |
| Very good | 0≤RSR≤0.5 | 0.75≤NSE≤1 | PBIAS<±10 | PBIAS<±15 | PBIAS<±25 |
| Good | 0.5≤RSR≤0.6 | 0.65≤NSE≤0.75 | ±10≤PBIAS<±15 | ±15≤PBIAS<±30 | ±25≤PBIAS<±40 |
| Satisfactory | 0.6≤RSR≤0.7 | 0.5≤RSR≤0.65 | ±15≤PBIAS<±25 | ±30≤PBIAS<±55 | ±40≤PBIAS<±70 |
| Unsatisfactory | RSR>0.7 | RSR>0.5 | PBIAS≥ ±25 | PBIAS≥ ±55 | PBIAS≥ ±70 |

Initial model results for daily flow yielded the following performance values: 1.18 for RSR, -0.39 for NSE, -42.87 for PBIAS and 0.05 for $R^2$ (determination coefficient). Monthly values of the performance statistics were: 0.61 for RSR, 0.60 for NSE, -45.75 for PBIAS and 0.71 for $R^2$. Thus, according to Moriasi et al. (2007) article, we can regard the results of the initial model run as satisfactory for both RSR and NSE and unsatisfactory for PBIAS.

Calibration of flow was performed by using the SUFI autocalibration algorithm in the SWAT-CUP2 2.1.5 program. This is a public domain program created by Karim Abbaspour and Raghvan Srinivasan (Eawag, 2008). 9 parameters were chosen for final flow calibration.

Initial selection of parameters was based on the results of sensitivity analysis obtained from the ArcSWAT program (sensitivity analysis module). Initially, 15 parameters were selected. Several runs with SWAT-CUP2 were made. After no more increase in the performance statistical criteria was observed, the initial autocalibration was stopped and each parameter's influence on the incremental increase in the performance statistical criteria was calculated. Based on this, four parameters were identified as most important for calibration. They were baseflow alpha factor (ALPHA_BF) expressed in days, groundwater delay time (GW_DELAY) expressed in days, maximum canopy storage (CANMX) expressed in mm $H_2O$ and effective hydraulic conductivity in main channel alluvium (CH_K2) expressed in mm/hr. The initial autocalibration run (100 iterations) for a daily time step with absolute possible values yielded the following performance values: 0.76 for RSR, 0.42 for NSE, -29.71 for PBIAS and 0.5 for $R^2$. For monthly run values the performance statistics were as follows: 0.57 for RSR, 0.65 for NSE, -29.35 for PBIAS and 0.72 for $R^2$. One by one other sensitive parameters were added to autocalibration, while letting other (already selected) parameters vary within SWAT-CUP2 suggested boundaries. In this way 5 additional parameters were added, which increased the performance statistics. They were snow pack temperature lag factor (TIMP), soil evaporation compensation factor (ESCO), threshold depth of water in the shallow aquifer required for return flow to occur (GWQMN) expressed in mm $H_2O$, surface runoff lag coefficient (SURLAG) and initial SCS runoff curve number for moisture condition II (CN2). This sequence was performed to minimize the number of parameters involved in autocalibration. The calibrated parameters as well as the initial values are presented in Table 6.

It is important to mention that during final steps[10] of calibration certain modifications were made

9 This coefficient provides the evaluation of the predictive power of hydrological models. It is calculated by dividing the sum of square root errors between modeled and observed results with the sum of square root errors between observed results and their mean and substacting final result from 1.

10 The final steps during calibration of nitrate loads. Introduced management practices slightly altered calibration values for flows. Therefore final calibrated model values represent not best autocalibration values for flows, but final prepared model flow values.

for the management practices. For BARL, CANP, SGBT land uses, tillage operations were introduced at 0.1 (generic spring plowing operation) and at 1.3 (generic fall plowing operation) fraction of the total heat units. For WWHT land use, a tillage operation was introduced just at 1.3 (generic fall plowing operation) fraction of the total heat units. An assumption was made that the main tillage operations occur just before planting and after harvesting of crops. Moreover, tile drainage management was activated for all land uses with the exception of FRSD and WATR. The depth to surface drainage (DDRAIN) has been placed at 1100 mm (value based on Smitiene (2007) thesis). Time to drain soil to field capacity (TDRAIN) was set to 24 hours and drain tile lag time (GDRAIN) to 48 hours. Those values were chosen based on the values adopted by similar studies.

*Table 6: Changes in parameter values during hydrology calibration.*

| | Initial value | Calibrated |
|---|---|---|
| ALPHA_BF (in .gw) | 1 | 0.5579 |
| GW_DELAY (in .gw) | 31 | 2.1048 |
| CH_K2 (in .rte) | 0 | 110.79 |
| CANMX (in .hru) | 0 | 35.0685 |
| TIMP (in .bsn) | 1 | 0.4785 |
| ESCO (in .hru) | 0 | 0.6189 |
| GWQMN (in .gw) | 0 | 426 |
| SURLAG (in .bsn) | 4 | 1.0107 |
| CN2 (in .mgt) | Dependent on land use | Multipified by 1.126 |

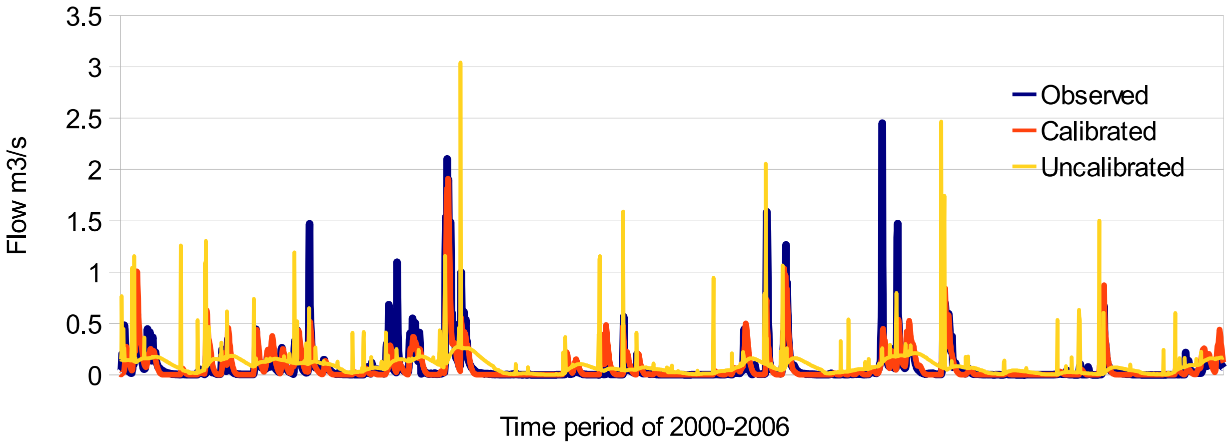
Final results of autocalibration yielded 0.63 for RSR, 0.63 for NSE, 0.62 for $R^2$ and -8.12 for PBIAS. For monthly run the performance statistics were: 0.45 for RSR, 0.78 for NSE, -9.21 for PBIAS and 0.81 for $R^2$. Thus, according to Moriasi et al. (2007) the model performance rating for calibration could be defined as very good for all statistical measures. The results of model performance statistics for the main calibration stages are presented in Table 7.

*Table 7: Results of the calibration of hydrology.*

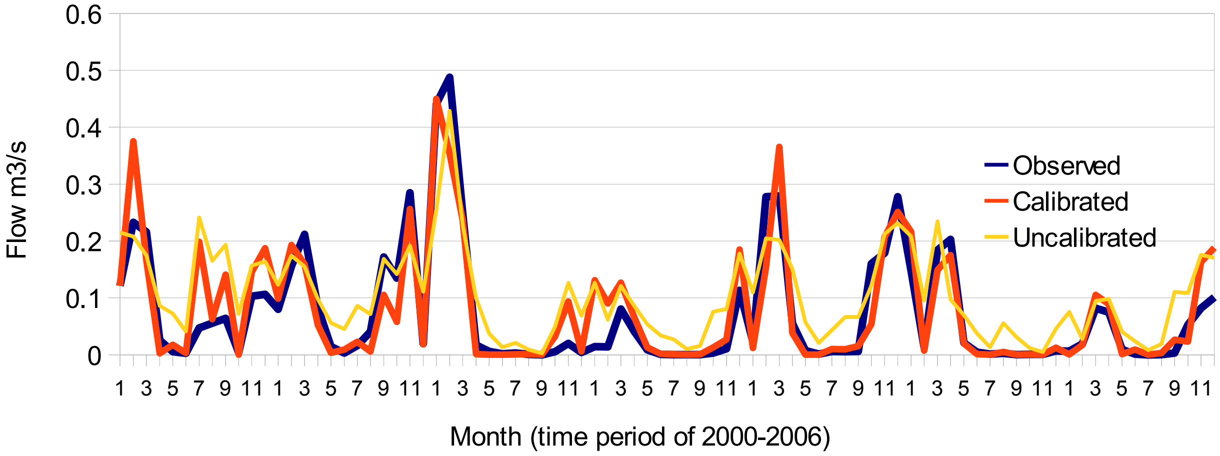
| | Before calibration | With 4 parameters | Calibrated |
|---|---|---|---|
| | Calibration for daily values | | |
| $R^2$ | 0.05 | 0.5 | 0.62 |
| RSR | 1.18 | 0.76 | 0.63 |
| NSE | -0.39 | 0.42 | 0.61 |
| PBIAS | -42.87 | -29.71 | -8.12 |
| | Calibration for monthly values | | |
| $R^2$ | 0.71 | 0.72 | 0.81 |
| RSR | 0.61 | 0.57 | 0.45 |
| NSE | 0.6 | 0.65 | 0.78 |
| PBIAS | -45.75 | -29.35 | -9.21 |

When analyzing flow results of the initial model run in comparison with the observation data (see Picture 8), it can be concluded, that the model default values quite well estimated the timing of peaks. However, overall surface water flows were overestimated and the base-flow level was too high. Moreover, the height of peaks was also wrong. Calibration allowed to get the base-flow right as well as the level of peaks. Furthermore, overall water balance had just a slight bias from the observation data. However, one problem was noticeable on the model run for daily values. A few highest peaks had not been simulated well by the model. Nevertheless, since this problem had little influence on monthly values, it was not considered important.

Initial results of the model run for monthly flow values produced quite good match (see Picture 9) between the measured and the simulated monthly flow average values ($R^2$ was 0.71 and NSE 0.6). However, the overall amount of water was highly overestimated (PBIAS -45.75). After calibration, the simulated average monthly flow values matched very good with the measured data. This statement is supported with the statistics provided in Table 7.



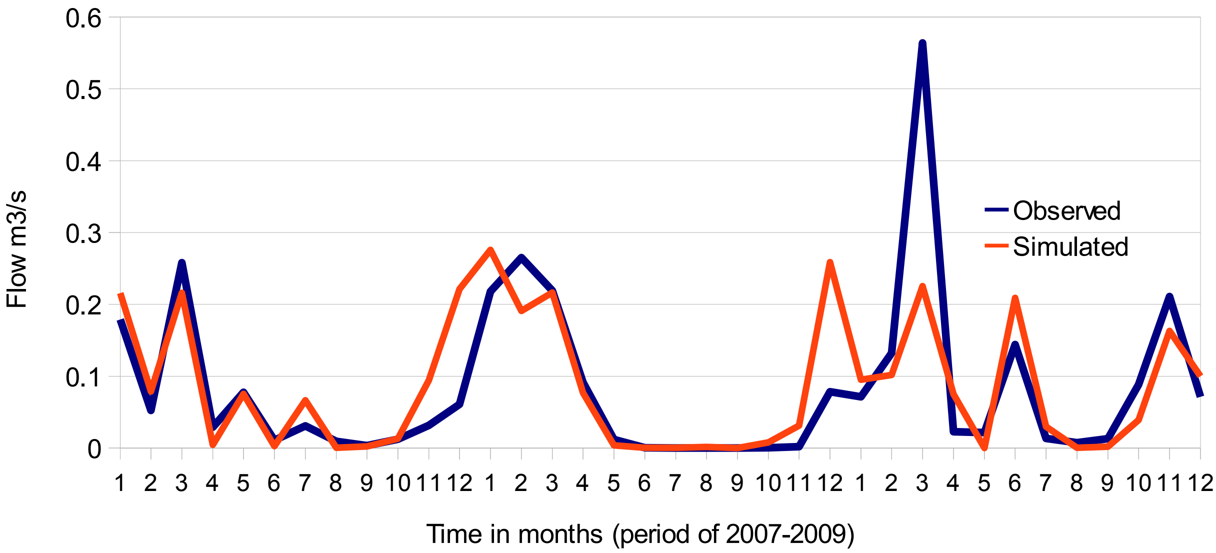

*Picture 8: Calibration for daily flow average values.*




*Picture 9: Calibration for monthly flow average values.*

### 5.4.3 Validation of Flow

The time period from the beginning of 2007 to the end of 2009 was selected for validation. A 5 year-long model warm-up period was necessary to obtain better results. During validation for flows the model performed worse than during calibration. The results of the model validation for monthly flow averages were 0.42 for RSR, 0.55 for NSE, 0.69 for $R^2$ and -3.15 for PBIAS. Peak flows before the beginning and during beginning of year 2009 were simulated incorrectly (see Picture 10). This might be also due to some changes in the flow measurement equipment or some natural changes in the environment, which had not been faced during calibration. Nevertheless, according to Moriasi et al. (2007), the model performance during validation could be rated as very good for RSR and PBIAS, while just satisfactory for NSE. Despite lower results during the validation of $R^2$ and NSE, the model could be used in regard to hydrology.




*Picture 10: Validation for monthly flow average values.*

### 5.4.4 Calibration of Nitrates

The Graisupis river catchment is dominated by agricultural activities and agricultural lands. As a consequence, nitrate concentrations do not meet the defined good water quality criteria. It is the key parameter responsible for the bad status of the water body. Table 8 provides yearly average values of all physical-chemical parameters required in the definition of the water body status (rules for the definition of water body status were issued by the order of 2010 of the minister of Environment of the Republic of Lithuania). Different colors in the table represent different water body statuses in regard to the analyzed parameter. Blue color represents a very good water body status, green – good, yellow – moderate, and orange - bad. It is clearly seen that nitrates and total nitrogen are the parameters responsible for the status of the Graisupis river. According to the water body status definition rules the overall water body status is in the category of the worst parameter[11]. However, high concentration of total nitrogen in the Graisupis river is driven mainly by high nitrate

11 If there are no biological parameters. Otherwise more complicated rules are applied. However with regard to watershed modeling this rule is the only one.

concentration[12], which constitutes around 83 % of total nitrogen concentration. Based on this information it is possible to narrow the assessment procedure by selecting nitrates as the only factor influencing water quality in this watershed. Thus, the assessment and abatement measures should be focused on the reduction of nitrate concentration in the Graisupis river.

*Table 8: Yearly averages of 7 water quality parameters.*

| Year | $NH_4$ | $N0_3$ | $N_{tot}$ | $PO_4$ | $P_{tot}$ | $BOD_7$ | $O_2$ |
|---|---|---|---|---|---|---|---|
| 1998 | 0.11 | 4.82 | 6.43 | 0.03 | 0.06 | -- | – |
| 1999 | 0.16 | 4.05 | 5.09 | 0.15 | 0.23 | 3.27 | – |
| 2000 | 0.09 | 5.67 | 6.55 | 0.06 | 0.09 | 1.7 | – |
| 2001 | 0.04 | 5.03 | 5.76 | 0.12 | 0.15 | 2.34 | – |
| 2002 | 0.08 | 3.2 | 3.92 | 0.15 | 0.29 | 3.49 | – |
| 2003 | 0.15 | 2.86 | 3.98 | 0.11 | 0.21 | 2.12 | – |
| 2004 | 0.15 | 6.05 | 6.75 | 0.07 | 0.11 | 1.8 | – |
| 2005 | 0.14 | 2.43 | 3.1 | 0.11 | 0.22 | 2.56 | 8.72 |
| 2006 | 0.19 | 9.88 | 11.18 | 0.14 | 0.22 | 4.77 | 7.14 |
| 2007 | 0.2 | 5.56 | 6.78 | 0.05 | 0.07 | 2.05 | – |
| 2008 | 0.06 | 4.92 | 6.51 | 0.07 | 0.16 | 3.28 | 10.17 |
| 2009 | 0.15 | 5.94 | 6.88 | 0.08 | 0.14 | 2.54 | 9.82 |

After calibration for flows the SWAT model was prepared for the simulation of nitrates. The inclusion of tillage operations was described in the Calibration of Flow section. Fertilization operations were altered as well (from default parameters). The fertilization operation was placed at 0.16 fraction of the total heat units. An amount of nitrogen applied[13] for the altered land use fertilization practices was the following: for WWHT - 111 kg/ha, for BARL – 48.5 kg/ha, for SGBT – 103 kg/ha, CANP – 70 kg/ha and for PAST – 74 kg/ha. The data was provided by Dr. Ausra Smitiene. It is based on the observations in the fields of Pikeliai and Lipliuniai located in the Graisupis river catchment (Smitiene, 2007). Moreover, nitrate concentration in the shallow aquifer (SHALLST_N) was introduced into the model. It was set to 2.3 mg N/l and was based on the WMI report of 2009.

Sensitivity analysis and autocalibration using the SWAT-CUP2 SUFI algorithm were performed for the selected parameters. Four parameters were used for nitrate calibration. They were: initial nitrate concentration in the soil layer (SOL_NO3), initial organic nitrogen concentration in the soil layer (SOL_ORGN), nitrogen percolation coefficient (NPERCO) and concentration of nitrogen in rainfall (RCN). Bounds used in autocalibration for RCN were based on the data bounds provided in Smitiene's dissertation (2007). Calibration for initial organic nitrogen and nitrate concentration in the soil was separated for upper and lower soil layers. Differences in the calibrated concentrations in upper and lower soil layer corresponded to the behavior of nitrogen compounds in natural systems. Upper layers due to plant residuals and humus have higher levels of nitrates and organic nitrogen. Changes in the parameters selected for nitrate load calibration are presented in Table 9.

*Table 9: Changes in parameter values during nitrate load calibration.*

| | Initial value | Calibrated |
|---|---|---|
| SOL_ORGN (1) (in .chm) | 0 | 6682.6 |
| SOL_ORGN (2) (in .chm) | 0 | 100 |
| SOL_NO3 (1) (in .chm) | 0 | 39.7 |

12 Concentrations of nitrite, amonium and organic nitrogen are quite low.

13 Elemental nitrogen was used from the SWAT fertilizer database.

| | | |
|---|---|---|
| SOL_NO3 (2) (in .chm) | 0 | 2.59 |
| NPERCO (in .bsn) | 0.2 | 0.416 |
| RCN (in bsn) | 1 | 6.94 |

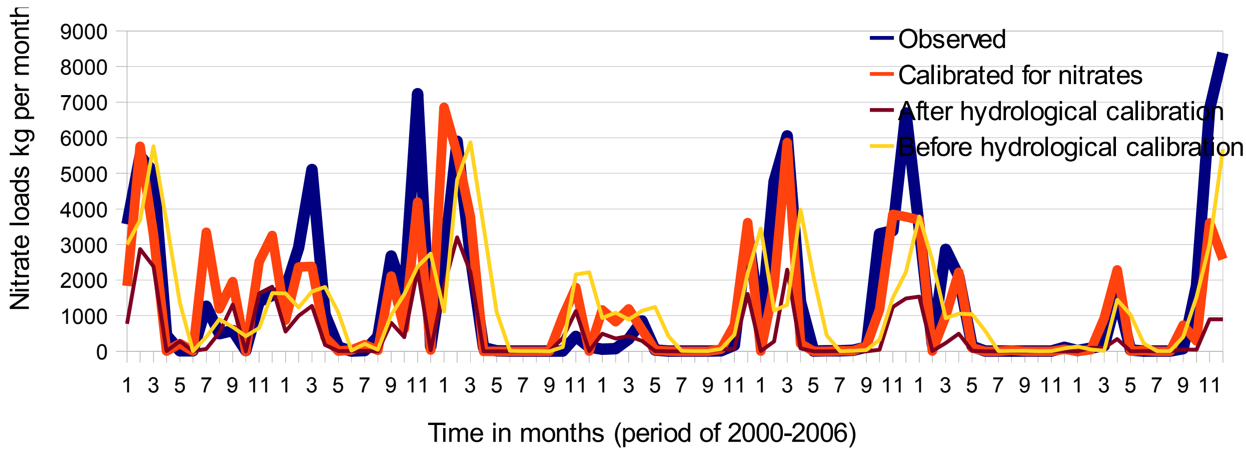
The model performance rating achieved during nitrate load calibration could be qualified as very good according to Moriasi et al. (2007). The PBIAS value for the calibrated model was equal to 13.73 (see Table 10). Moriasi et al. (2007) proposed a general calibration procedure for flow, sediment, and nutrients of the watershed model. According to this procedure, model calibration target for mineral nitrogen loads should be ±25 for PBIAS, 0.6 for RSR and 0.65 for NSE. The calibrated model was just slightly below in NSE coefficient. It is important to point out that nitrate loads were quite well simulated even with the uncalibrated model. As seen from the calibration results (see Table 10) hydrological calibration worsen the statistics for match between the simulated nitrate loads and the monitoring data. This pattern is stronger in the statistics for nitrate concentration. Therefore, it is important to note that probably the best way to calibrate the SWAT model could be when all parameters (hydrological and water quality) are simultaneously addressed.

*Table 10: Results of the calibration of nitrate loads and concentration simulation.*

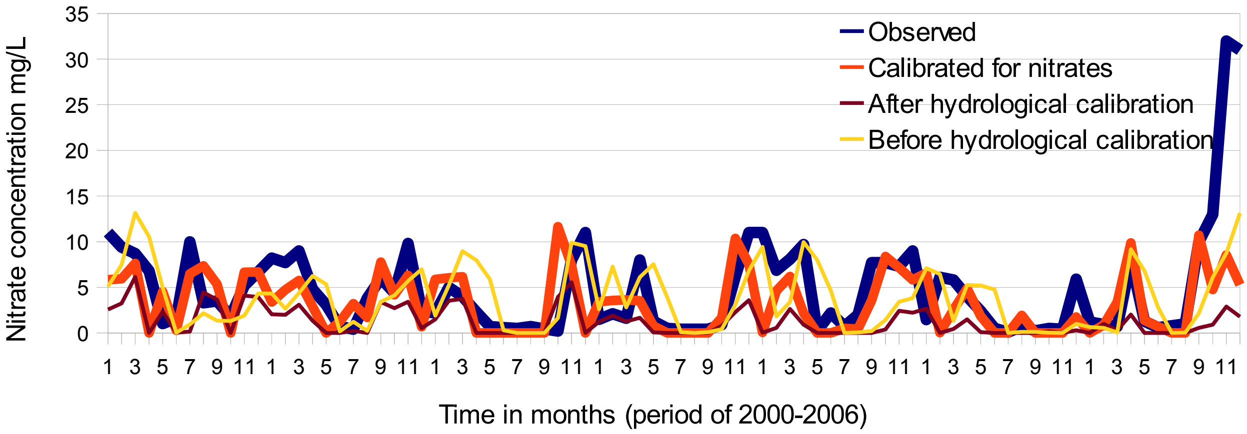
| | Before hydrological calibration | After hydrological calibration | Calibrated for nitrate |
|---|---|---|---|
| | Calibration for monthly loads | | |
| $R^2$ | 0.37 | 0.56 | 0.63 |
| NSE | 0.36 | 0.23 | 0.62 |
| PBIAS | 6.01 | 64.62 | 13.73 |
| | Results for monthly average concentrations | | |
| $R^2$ | 0.3 | 0.13 | 0.23 |
| NSE | 0.26 | -0.28 | 0.15 |
| PBIAS | 20.59 | 73.93 | 31.51 |

Comparing the simulation with the monitoring data on the time series graph (see Picture 11) it is possible to conclude that the model well matches monitoring data with the exception of a few peaks. It is hard to point out, what is the cause of these minor problems. Further analysis is be needed to provide an answer to this question. Nevertheless, despite these minor shortcomings the model could be used for nitrate load simulation.

Since the model was calibrated for nitrate loads, its results for the concentration simulation are not as satisfactory. It is important to note that the results were better before the calibration for water flows and nitrate loads than after. The model performance results for the simulation of nitrate concentration are provided in Table 10. From the time series graph (see Picture 12) it can be concluded that the model is good enough for the depiction of the variation in nitrate concentration. The simulation values varied according to the monitoring data (and in the abounds of the monitored data) with the exception of the end of the calibration period, which had unusually high concentration of nitrate for the monitoring data. This might be due to some additional anthropogenic factors not accounted by the model. Another explanation for these high concentrations is droughts, which were persistent during the summer and early autumn. The drought period was followed by intensive rains. Thus nitrates, which had been accumulated during droughts, were simultaneously washed during the rainy period. This explanation was supplied in the WMI report of 2007.




*Picture 11: Calibration for monthly nitrate loads for the Graisupis river catchment.*




*Picture 12: Simulation of nitrate concentration for calibration period compared with monitoring data.*

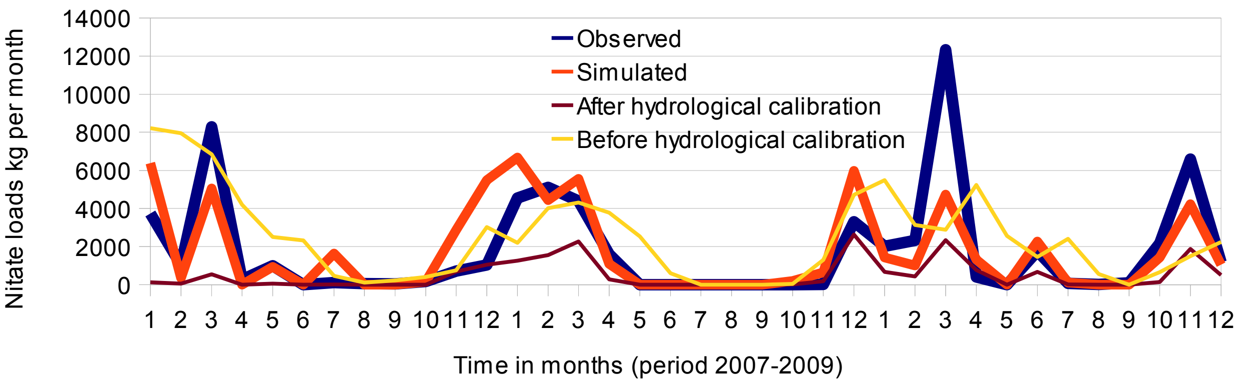
## 5.4.5 Validation of Nitrates

The model performance results for the validation period are presented in Table 11. Those results are lower (in regard to model performance) than obtained for the calibration period for $R^2$ and NSE, however, higher than for PBIAS statistics. Nevertheless, if taking into regard the results of the water flow simulation during validation (which were noticeably worse compared to calibration results), it would be possible to state that simulation of nitrogen processes by the model was relatively better during validation. It is well seen in $R^2$ statistics. For instance, the difference in $R^2$ between the model performance for monthly water flow and the monthly nitrate loads simulation was 0.18 for the calibration period. Yet, for the validation period, this difference was only 0.02. An improvement in nitrate simulation could be further emphasized by pointing out to the improved concentration simulation statistics.

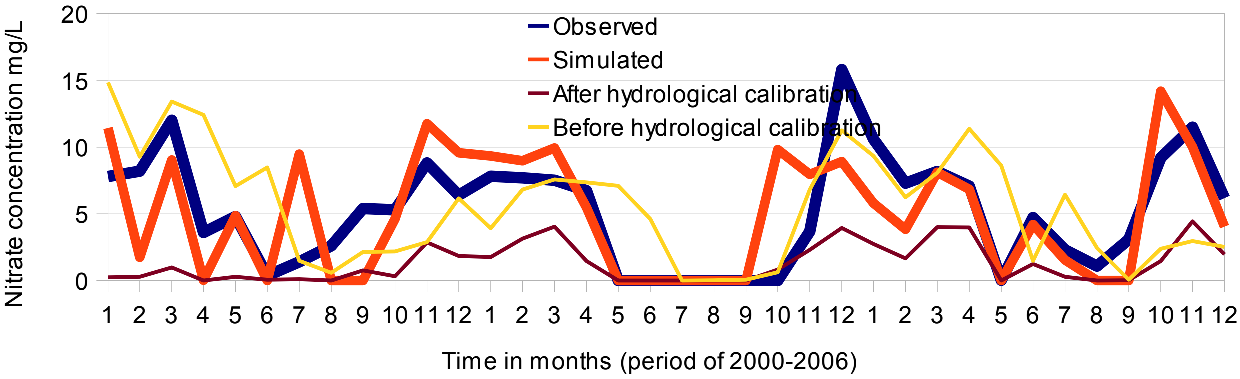
*Table 11: Results of the validation of nitrate loads and concentration simulation.*

| | Before hydrological calibration | After hydrological calibration | Calibrated for nitrate |
|---|---|---|---|
| | Validation for monthly loads | | |
| R2 | 0.16 | 0.53 | 0.53 |
| NSE | -0.08 | 0.1 | 0.53 |
| PBIAS | -35.54 | 71.73 | -0.92 |
| | Results for monthly average concentrations for validation period | | |
| R2 | 0.21 | 0.52 | 0.43 |
| NSE | -0.11 | -0.55 | 0.23 |
| PBIAS | -6.32 | 74.9 | 3.04 |

The time series graph (see Picture 13) shows adequate model behaviors with the exception of one peek load for the validation period for nitrate load simulation. The same time step was problematic in the concentration simulation (see Picture 14). Yet, the difference between the observed and simulated nitrate concentration values is more significant than for nitrate load simulation. This might be explained by the model's calibration only for nitrate loads.




*Picture 13: Validation for monthly nitrate loads for the Graisupis river catchment.*




*Picture 14: Simulation of nitrate concentration for validation period.*

## 5.5 CSAs and Sensitive Areas Identification

The identification of CSAs was made by using nitrate loading results from the prepared model. The SWAT model was run for all 12 year period from 1998 to 2009 with a 5 year-long warm-up period. Yearly values were saved and then averaged for 12 year period. NSURQ[14], NLATQ[15] and NO3GW[16] output variables from output.hru file were added together as they represented total nitrate loading to surface water bodies from separate HRUs. Thus, nitrate contribution[17] was identified for all areas in the watershed. CSAs was defined in a way, which was proposed in the article of Sivertun & Prange (2003), i.e., based on the deviation of nitrate loading from average values, areas were identified as risk or sub-risk areas. This method was chosen, because of its simplicity. It would have been much more complicated to provide some threshold (of tolerated nitrate loading) if the country context was not considered.

The identification of sensitive areas was based on the assumption that area's sensitivity is related to the HRU's response to nitrogen loads. A more sensitive area would contribute more nitrates into reaches compared to the less sensitive one, if the same amount of nitrate loads (for instance as a fertilizer) was applied on it. One important point should be stressed that land use should not be incorporated into sensitivity as it already represents the anthropogenic load. For example, forest, arable land or waste site would have quite different loading to reaches and quite different responses to additional load not due to different physical characteristics of the land (although it is possible as well), but due to the influence of difference in anthropogenic activities. For the identification of sensitive areas it is most important to find out the differences in response, which depend on physical characteristics of the land. Thus, for the identification of sensitive areas the land use factor should be eliminated. This was done by assigning pastures to the whole watershed area with the exception of water bodies. Soil and slope were left intact. The prepared model was run and nitrate loadings obtained as in the CSAs identification. In order to get a HRU's response, 100 kg/ha of elementary nitrogen as a fertilizer was added on the whole watershed area (except for the water bodies) at 0.16 of Heat Units. The model was run and the results of nitrate loads to surface water bodies were obtained. The difference of nitrate loads between the model runs was calculated for each HRU. Sensitivity of each HRU was calculated with Equation 3.

$$Sensitivity\ for\ HRU = \frac{Nitrate\ loading\ diffence\ for\ selected\ HRU - Minimal\ difference\ for\ all\ HRUs}{Maximal\ difference\ for\ all\ HRUs - Minimal\ difference\ for\ all\ HRUs}$$

*Equation 3: Formula for sensitivity calculation.*

The obtained results were used to produce a map of sensitivity to nitrogen load for the watershed. According to the sensitivity values, 5 arbitrary categories were established. Very low sensitivity area was assigned to the values from 0 to 0.2, low sensitivity – values from 0.2 to 0.4, medium sensitivity – values from 0.4 to 0.6, high sensitivity – values from 0.6 to 0.8 and very high sensitivity – values from 0.8 to 1.

14 $NO_3$ in surface runoff (kg N/ha). Nitrate transported with surface runoff into the reach during the time step.
15 $NO_3$ in lateral flow (kg N/ha). Nitrate transported by lateral flow into the reach during the time step.
16 $NO_3$ transported into main channel in the groundwater loading from the HRU (kg N/ha).
17 It was expressed in kg of nitrogen per hectare per year.

## 5.6 GA Preparation and Spatial Optimization

GA was used in this study for multi-objective spatial optimization purposes. It was built by using Visual Basic for Applications integrated into the Microsoft Office 2003 Excel software.

Before the preparation of GA, the BEP database was compiled. It consisted of nitrate loading and cost information for each HRU under separate scenarios. Scenarios represented the application of a certain BEP on HRU as well as doing nothing for the baseline scenario. Based on the applicability to the Graisupis river watershed, five BEPs were selected for assessment. They were: filter strips of 20 meters, cover crops, residue management, converting to grassland or afforestation. Nitrate loading to surface water bodies for the baseline scenario was obtained identically as in the CSA identification and the costs for all HRUs were assigned to 0. The prepared model was modified for the assessment of the BEPs' influence on nitrate loadings. As a guidance for the representation of BEPs representation Arabi et al. (2007) article of “Representation of agricultural conservation practices with SWAT” was used.

For the representation of filter strips, FILTERW[18] parameter was set to 20 meters for land uses of CORN, WWTH, CANP, BARL, SCRN, SGBT, PEAS, AGRC, AGRL and PAST. For cost calculation some assumptions were made. The area of filter strips for each HRU was approximated by multiplying the HRU area with the ratio between the sum of 20 meters buffer area around the streams (which was 0.77 $km^2$) and the watershed area (which was 14.2 $km^2$). All cost values were obtained from the LEPA's Nemunas river basin district management plan (2010). For all the mentioned land use types with the exception of PAST the cost was 1225 LTL[19] per hectare per year for operation and management and 1047 LTL for installation. For the PAST land use type it was 1288 LTL per hectare per year for operation and management and 100 LTL for installation.

For the representation of cover crops the method suggested in the article of Arabi et al. (2007) was used. The sequence of operations required to represent cover crops is presented in Table 12. According to the LEPA (2010) cover crops have no cost for installation. Operation and management costs are 350 LTL per hectare per year. Cover crops were applied to CORN, WWTH, CANP, BARL, SCRN, SGBT, PEAS, AGRC and AGRL land uses.

*Table 12: Representation of cover crops.*

| Year | Operation | Crop | Date | |
|---|---|---|---|---|
| | | | month | day |
| 1 | Plant begin | WWTH | March | 1 |
| 1 | Harvest and kill | WWTH | May | 2 |
| 1 | Tillage | | May | 3 |
| 1 | N-fertilizer | | May | 5 |
| 1 | Plant begin | Plant for HRU | May | 10 |
| | Harvest and kill | Plant for HRU | October | 15 |
| 1 | Plant begin | WWTH | October | 15 |
| 1 | Harvest and kill | WWTH | December | 31 |

18 Width of edge-of-field filter strip (m).
19 Litas exchange rate to euro is 3.45 to 1.

The representation of residue management was prepared according to the instructions provided in the article of Arabi et al. (2007). The altered parameters (for residue biomass left on surface of 500 kg/ha) are presented in Table 13. The cost of this BEP is assumed to be similar to the cost of organic farming. It is 700 LTL yearly for operation and management.

*Table 13: Parameters for representation of residue management.*

| Land use type | Initial USLE_C[20] | Modified USLE_C | Modified USLE_P[21] |
|---|---|---|---|
| WWTH | 0.03 | 0.13 | 0.24 |
| CORN | 0.2 | 0.37 | 0.55 |
| CANP | 0.2 | 0.37 | 0.55 |
| BARL | 0.01 | 0.07 | 0.15 |
| SCRN | 0.2 | 0.37 | 0.55 |
| PEAS | 0.2 | 0.37 | 0.55 |
| SGBT | 0.2 | 0.37 | 0.55 |
| AGRC | 0.03 | 0.13 | 0.24 |
| AGRL | 0.2 | 0.37 | 0.55 |

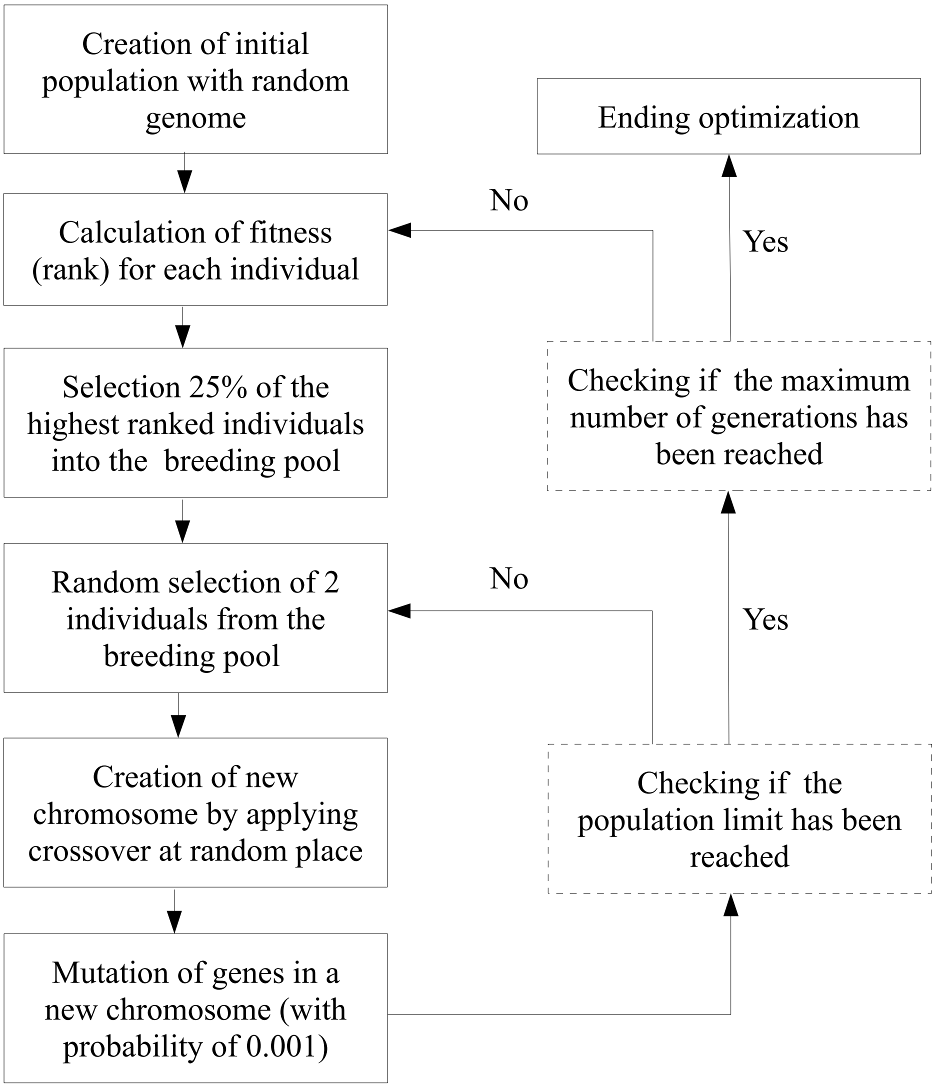
The BEPs of converting to grasslands and afforestation were represented by changing land uses to PAST or FRSD depending on which BEP was evaluated. The conversion to grasslands was applied on CORN, WWTH, CANP, BARL, SCRN, SGBT, PEAS, AGRC and AGRL land uses, whereas afforestation on CORN, WWTH, CANP, BARL, SCRN, SGBT, PEAS, AGRC, AGRL and PAST land uses. The cost of converting to grasslands given by the LEPA (2010) was 550 LTL per hectare per year on operation and management (no cost given for installation). Afforestation costs were 20000 LTL per hectare.

After the preparation of the BEPs database, the GA algorithm in Microsoft Visual Basic was prepared. The GA terminology is closely connected to the biological terminology. In order to understand the application of the GA for watershed problems, it is important to relate the GA terminology to the problem at hand. The Graisupis river catchment had 604 HRUs, each of which had at the most 6 different management options available: the baseline, filter strips, cover crops, residue management, conversion to grasslands, and afforestation. The combination of management options for all 604 HRUs was used as the main object for the optimization. It can be called the combined watershed scenario. 100 of these combined watershed scenarios were used in the each loop of the GA algorithm. This would translate into the GA "language" as following. The combined watershed scenario is equal to individual and information about the combined watershed scenario (or individual) is written into chromosome. Since prepared model for the Graisupis river catchment had 604 HRUs, chromosome in GA had 604 genes, for which at the most 6 alleles (6 management options) were available. The number of alleles depends on the applicability of BEPs on the HRU. For instance, the HRU with the land use type of FRSD (deciduous forest) had just 1 allele (baseline), while the HRU with CORN (corn) had all 6 (all management options). The rest links between GA and watershed modeling terminology are: selection of individuals would translate into selection of combined watershed scenarios based on their cost-effectiveness, mutation of gene - into

20 Minimum value of USLE_C factor for water erosion applicable to the land cover/plant.
21 USLE equation support factor.

randomly changing management option for HRU, crossover – merging parts of two combined watershed scenarios into one, generation number – number of loops GA would run.

The prepared GA algorithm used following steps in the optimization. Firstly, it randomly created the initial population of 100 individuals. For each individual a cumulative rank was calculated. It was done by averaging the individual's rank based on nitrate loading and the rank based on costs. Then, based on this cumulative rank value, individuals were ranked again. 25% of the highest ranked individuals were selected for reproduction. Out of this group two individuals were randomly selected, and by applying crossover technique (when parts of parent chromosomes are used in building an offspring chromosome), new individuals were created. The procedure was continued until the population reached 100 individuals again. Mutation was applied at the probability of 0.001 for gene per generation. This rate of mutation was selected after a few initial tests indicated its suitability in global optimum search. The described algorithm was looped the number of generations used in GA. The framework of developed algorithm is presented in Picture 15 and the code of algorithm implemented in Microsoft Visual Basic is presented in Appendix B.



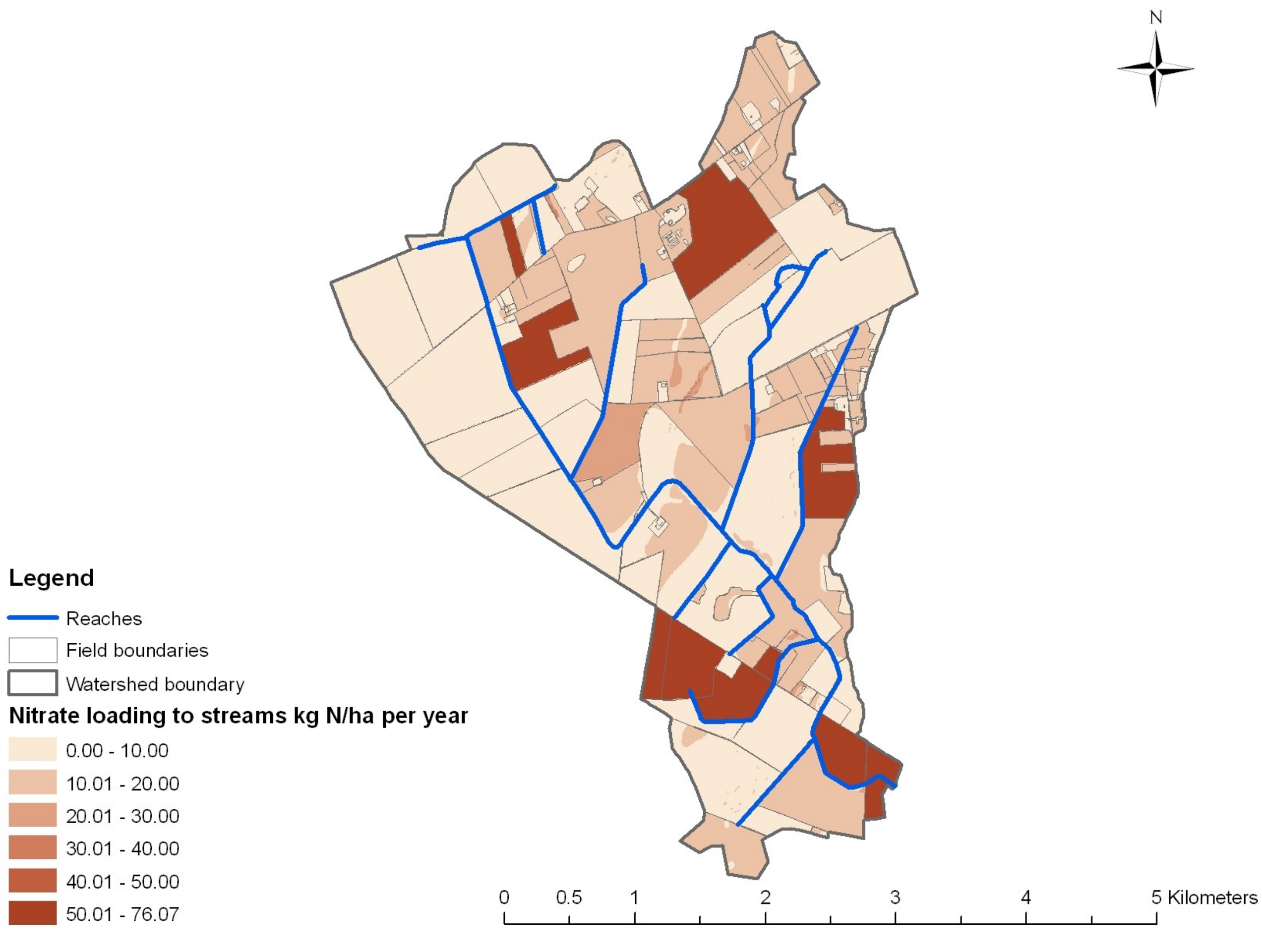

*Picture 15: GA optimization algorithm.*

# 6 Results

The results of the analysis CSAs, sensitive areas and multi-objective spatial optimization are presented in this section.

## *6.1 CSAs and Sensitive Areas*

Nitrate loading was calculated before the identification of CSAs. The map of nitrate loading (see Picture 16) allows to see the area's contribution to pollution loading in the Graisupis river watershed. 12% of the territory produces nitrate loading to surface water bodies at rate of more than 50 kg N/ha per year. Those areas in the map are represented with the darkest color. Such areas are dominated with WWTH land use. The large contribution of loadings from these areas could be explained with high fertilization rates (111 kg N per hectare per year) of these fields. Other large contributions of nitrate loadings come from AGRC and BARL land use (30 to 40 kg N/ha per year).

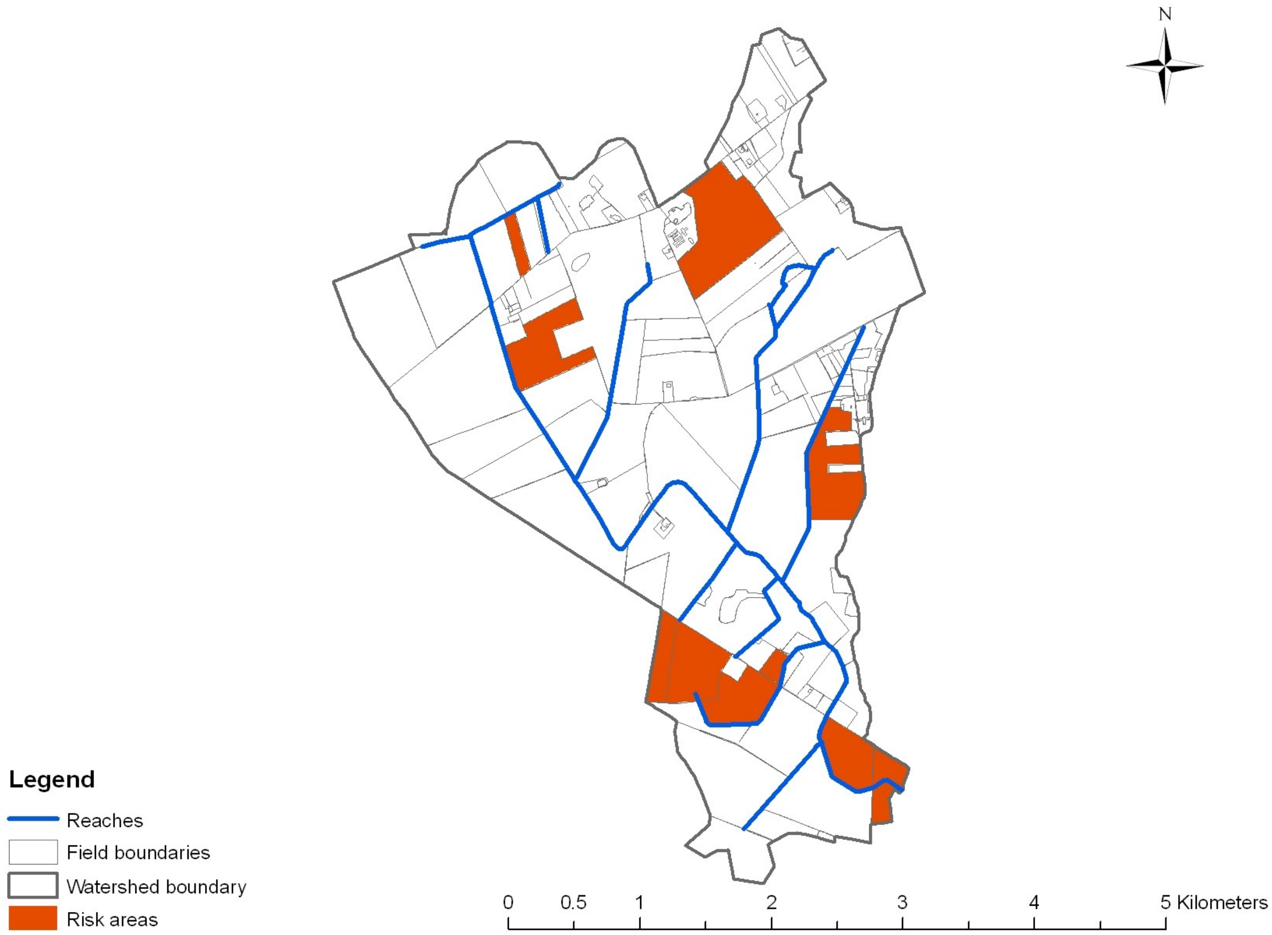



*Picture 16: Nitrate loading from different areas in the Graisupis river catchment.*

*Table 14: Number of HRUs, area occupied and percentage from whole watershed divided between different nitrate loading categories.*

| | 0-10 kg N/ha | 10-20 kg N/ha | 20-30 kg N/ha | 30-40 kg N/ha | 50-77 kg N/ha |
|---|---|---|---|---|---|
| Nb. of HRUs | 325 | 180 | 29 | 0 | 67 |
| Area in ha | 760 | 452 | 27 | 0 | 175 |
| % of watershed | 54 | 32 | 2 | 0 | 12 |

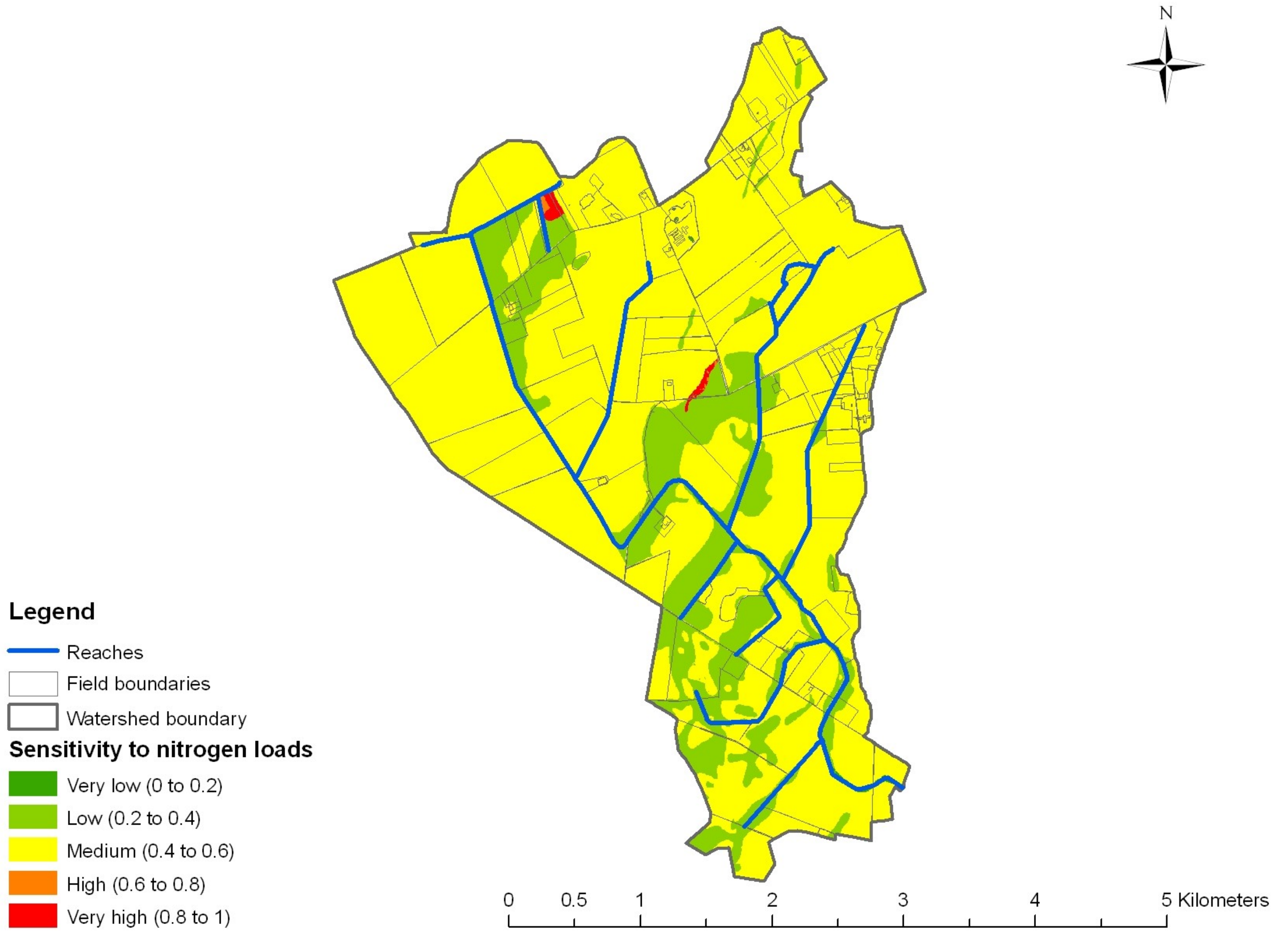
For the identification of CSAs the method proposed by Sivertun & Prange (2003) was used. CSAs were identified as risk and sub-risk areas based on the deviation of nitrate loading from average values. Average nitrate loading was 15.9 kg N/ha per year. Sub-risk areas were defined as areas, which deviated from the average value by one to two standard deviations. Sub-risk areas enclosed areas with 34.9-54.5 kg N/ha of nitrate loading per year. However, the Graisupis river watershed had no such areas falling into the mentioned interval of nitrate loading. Therefore, only risk areas were identified, which were defined by the deviation of more than two standard deviations from average values. It is the interval of more than 54.5 kg N/ha per year. The map of risk areas is presented in Picture 17. Risk areas occupy 1.75 square kilometers. This accounts for 12.4% of the Graisupis river catchment.




*Picture 17: CSAs for the Graisupis river catchment presented as risk areas.*

The identification of sensitive areas to nitrogen loads is important for land management purposes. Authorities responsible for watershed management should be able to prioritize areas based on their physical characteristics for nutrient or pollutant load retention. Areas, which have higher retention abilities, could be used for more intensive agricultural activities compared to the areas, which have less retention abilities. Thus, sensitive areas might be subject to the limitation of fertilizer application or other types of BEPs.

The results of the sensitive areas analysis are presented in Picture 18 and Table 15. Based on the used methodology, most of the areas in the Graisupis river catchment fall in the group of medium sensitivity. It should be emphasized that the way of assigning sensitivity values are quite arbitrary and it would be hard to compare with other watersheds. To get a better understanding of sensitivity, the whole country context should be considered. In other words, it should be defined on a country basis what is meant by sensitive and what is meant by not sensitive. Now it is done by looking into the maximum and minimum changes of land responses to nitrogen fertilization.




*Picture 18: Spatial distribution of sensitivity to nitrogen loads in the Graisupis river watershed.*

Medium and low sensitivity areas occupy 99.5% of the watershed, while rest three categories occupy only around 0.6%. It is hard to compare results with other studies, since no similar studies has been observed in literature[22]. However results could be compared with general patterns of non-

22 Literature review on the identification of sensitivity was not extensive due to scope of the study.

point pollution, which usually follow log-normal distribution (Diebel et al., 2008). According to statistical values of distribution fitting (p-value from Kolmogorov-Smirnov test was less than 0.01 and Chi-Square less than 0.0000001) distribution of sensitivity results can be approximated with log-normal distribution. Thus results are consistent with general non-point pollution patterns.

High and very high sensitivity values were obtained in the areas, which were covered with Gleyic Arenosols and Haplic Arenosols soil type. The lowest sensitivity values were mainly in the areas dominated by Calcaric Arenosols soil type. Results suggest that it is possible to identify the sensitivity of areas based only on soil variables. This is due to elimination of land-use factor, little change in the elevation of such small area and no changes in climate variables for the catchment area (data only from one meteorological station). However this will not be possible if method would be applied on larger watersheds, where are more variation in elevation and data from multiple meteorological stations are used.

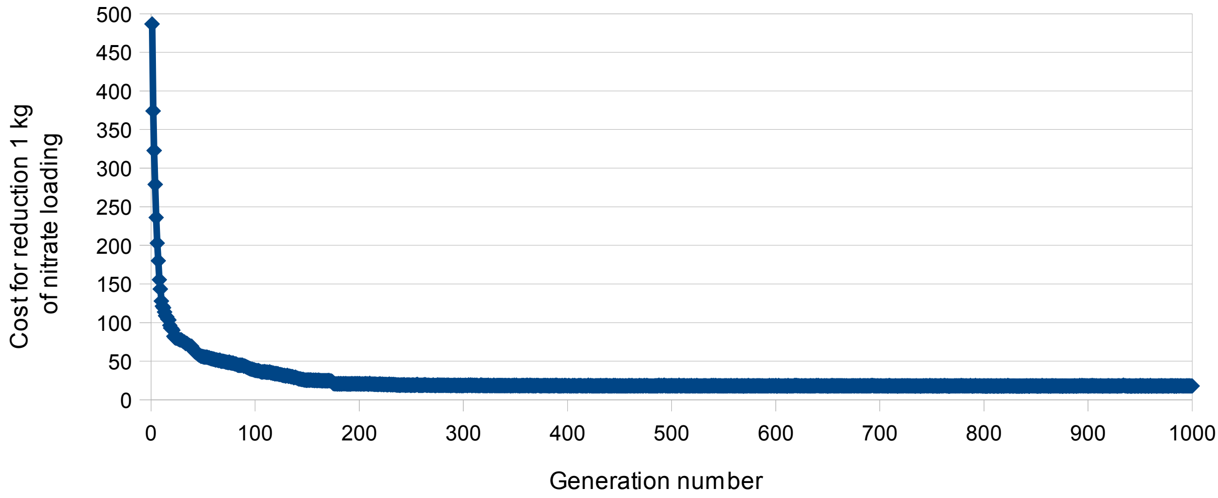
*Table 15: Number of HRUs, area occupied and percentage from the whole watershed divided between different sensitivity categories.*

| | Very low | Low | Medium | High | Very high |
|---|---|---|---|---|---|
| Nb. of HRUs | 7 | 242 | 338 | 8 | 9 |
| Area in ha | 2 | 253 | 1155 | 1 | 4 |
| % of watershed | 0.2 | 17.9 | 81.6 | 0.1 | 0.3 |

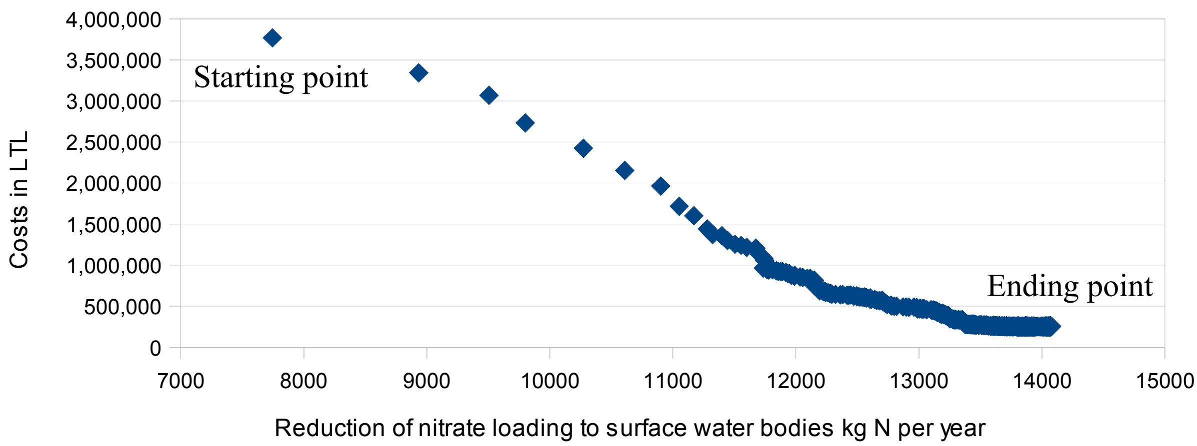
## 6.2 Multi-objective Spatial Optimization

Two types of objectives were raised for MOSO. Firstly, to find out the most cost-effective option, which would balance the environmental benefits with the economic costs. Secondly, to depict Pareto optimum, which would provide information for stakeholders about the trade-offs between environmental and economic objectives. This information would allow stakeholders to choose by themselves what environmental aims are the most appropriate or reachable (based on available funds) for them.

For the first objective, 1000 generations in GA were used. The evolution of cost-effectiveness for the reduction of nitrate loading to surface water bodies is presented in Picture 19 and in Picture 20. Under the baseline scenario, the total nitrate loading to surface water bodies in the Graisupis river watershed was 22487 kg N per year. Loadings of nitrate were decreased by nearly 8000 kg (reducing the total loading by 34%) when the scenarios (baseline including) were selected randomly for each HRU. However, the cost of the random scenario was around 3,800,000 LTL. That is, the cost for reducing 1 kg N of nitrate loading would be around 480 LTL. By applying GA optimization within 200 generations it was possible to increase the cost-effectiveness of pollution reduction up to 24 times. The most cost-effective solution, when running GA optimization for 1000 generations, yielded in 16.9 LTL for reduction of 1 kg N of nitrate loading to surface water bodies. This solution also reduced nitrate loading by 13,967 kg N per year (62% reduction from the total) and with only 236,605 LTL (6% of the cost for the random scenario).

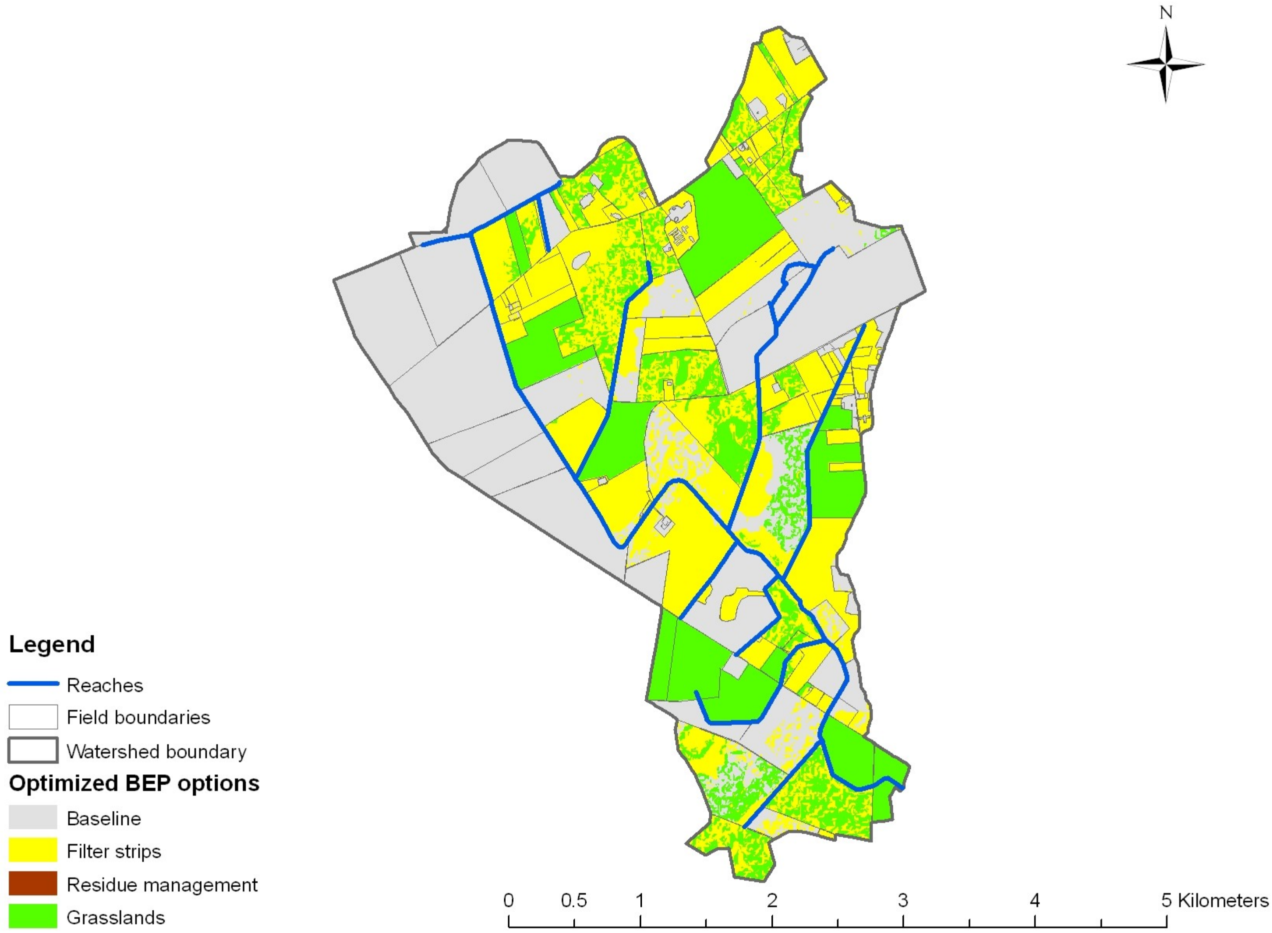



*Picture 19: Optimization of costs of reduction of nitrate loading to surface water bodies depending on generation.*




*Picture 20: Progress of cost reduction and reduction of nitrate loading during optimization.*

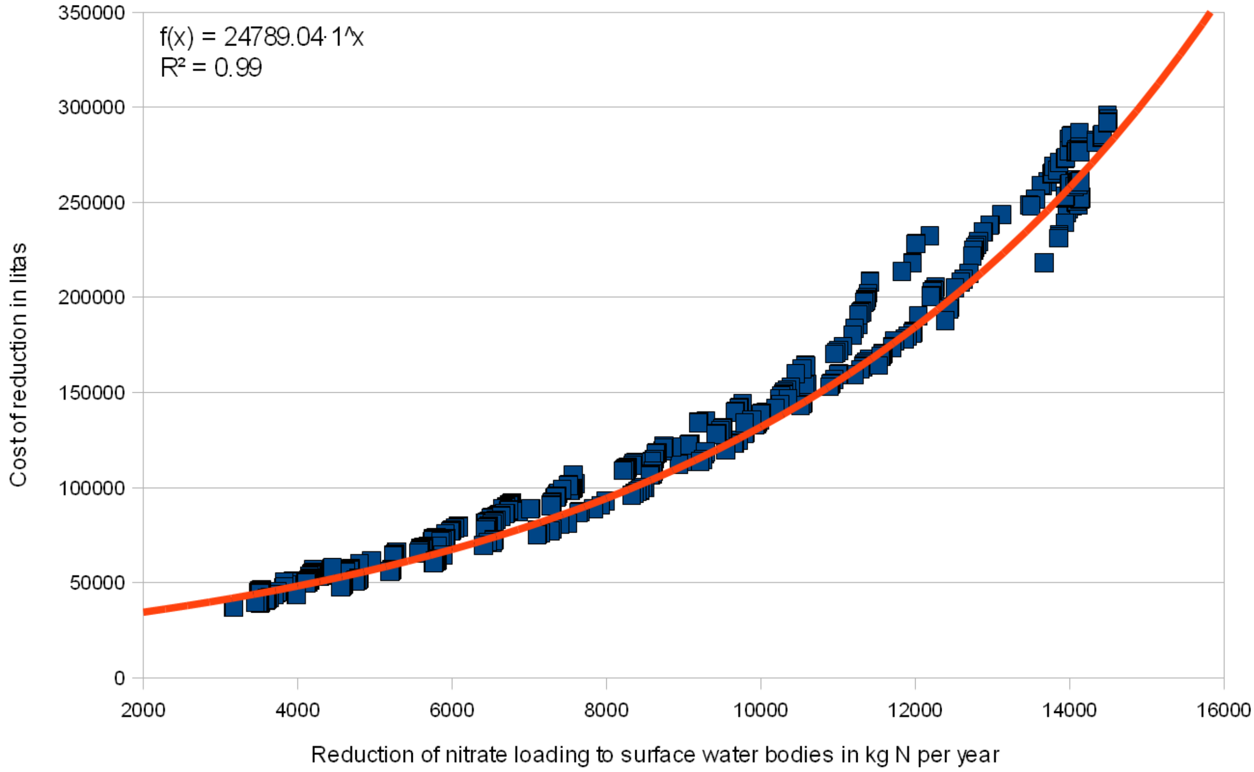
The previously mentioned most cost-effective scenario for the Graisupis river catchment is presented in Picture 21. It is well seen that filter strips and conversion to grasslands are the most cost-effective BEPs for the watershed. There is also the BEP of residue management, yet the area for it is very small. According to this scenario, the baseline should occupy 42% of the watershed, grasslands – 24%, residue management - 0.001% and filter strips should be considered for 34% of the catchment territory. Grasslands are suggested for the most critical areas.




*Picture 21: The most cost-effective scenarion reached for the first objective.*

In order to reach the second objective, the GA optimization algorithm was modified. It was divided into three steps. Firstly, the GA optimization algorithm was run for 1000 generations. When this number of generations was reached the whole population was saved separately. For the next 1000 generations only those solutions, which had higher the loads than loads average for the previous generation, were selected for mating. This allowed the optimization algorithm to travel through the Pareto optimum front in one direction. Then, after 1000 generations saved population was used to come back to the same point and travel to another direction. Therefore, 3000 generations were used for getting Pareto optimum: the first thousand for coming to Pareto optimum, the second thousand for traveling half of the front and the third thousand for traveling the rest of the front. GA optimization was repeated 3 times in order to obtain better dispersion along Pareto optimum. Thus, in total 9000 generations were used. The results of this part of optimization are presented in Picture 22[23].

23 Picture shows just one tenth of the total number of sollution points due to the problems with chart stability. The Pareto optimum front should be much thicker.

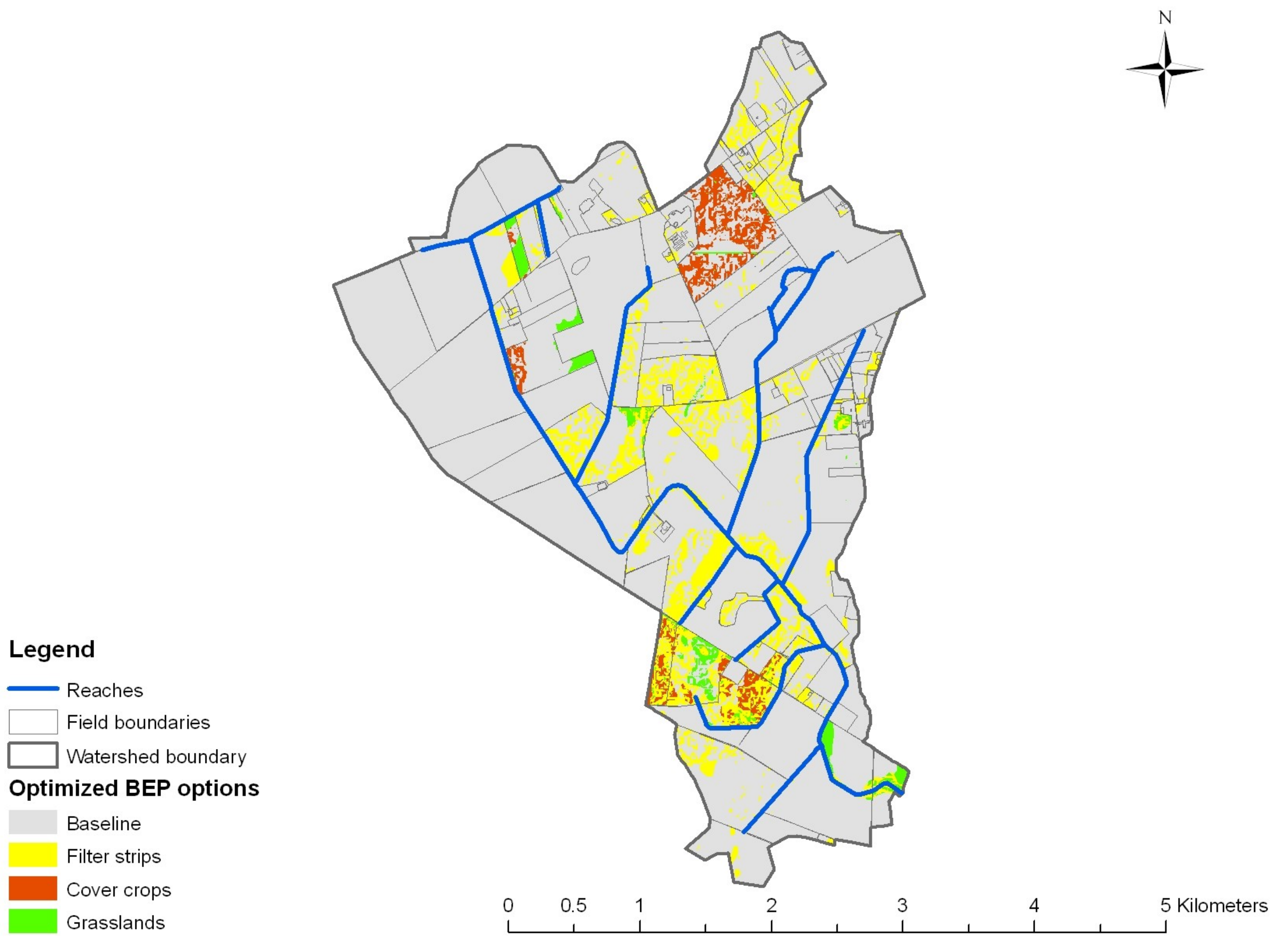



*Picture 22: Trade-offs between reduction of nitrate loading to surface water bodies and its cost.*

Picture 22 provides a good representation of the trade-offs between economic and environmental factors. Around 300,000 LTL should be spent if the maximum reduction of nitrate loading to surface water bodies is aimed at. However, if the aim is not as high, the cost of reduction could go down to 50,000 LTL.

The most cost-effective scenario through all Pareto optimum front was 9.7 LTL for reduction of 1 kg N of nitrate loading. For this scenario, the reduction of 5719 kg N of nitrate loading could be achieved only at the cost of 55,568 LTL. If compared to the previous cost-effective scenario, those numbers could be translated into: 41% of the previous reduction might be obtained with 23% of the previous costs.

The most cost-effective solution from the Pareto optimum front is presented in Picture 23. Areas, which are not under the baseline scenario, are much smaller compared to the solution reached for the first objective. Grasslands and filter strips are also the most cost-effective BEPs. Yet cover crops come to this list as well. According to this solution, the baseline should occupy 85% of the watershed, cover crops - 2.5%, grasslands - 1.6% and filter strips should be considered for 11% of the territory.




*Picture 23: The most cost-effective scenario from Pareto optimum.*

# 7 Discussion

It is important to discuss a few issues before going to the conclusions of this work. Understanding is necessary about the limitations of this work as well as future directions for research. Moreover, it is also essential to discuss this study's value, possible benefits, social aspects involved and relation to the science of geography.

The major problem during the model preparation was lack of data. Especially, the lack of soil parameters and land management data (fertilization, application of pesticides, tillage operations, sowing and harvesting). Since the Graisupis river catchment is the study area of the WMI, more data here were available here comparing to the rest of the country. However, looking from the diffuse pollution modeling perspective, certain corrections in monitoring activities should be made in order to obtain the missing data. The missing parameters were estimated with different methods or assumptions, or gathered from databases prepared for the world.

In general the quality and consistency of available data (water flow, water quality, weather data, etc) were quite good. Monitoring activities in the area were continuous since year 1998. One of the problems, which has influenced the confidence level of modeling results, was the location of Dotnuva meteorological station (the primary source of weather data). It was located about 4 kilometers away from the catchment and might not very well represented precipitation patterns in the catchment. There was also problem with the rain gauges, which had not been designed to shield wind effects, thus increasing uncertainties in the precipitation data. Furthermore, the solar radiation data were collected about 47 km away from the catchment. This can be the reason behind the modeling problems of daily flow in regard to some peaks. DEM, although was good enough for the study objectives, had minor problems. In some places, where sheets were touching each other (all PEM was provided in sheets covering area of 5 km X 5 km), there was mismatch of elevation up to 0.5 meters, which resulted in some errors during slope definition. This problem had little influence on the results of this study, however one have to be aware of it, if the same database should be used to prepare the SWAT model on different watersheds. Furthermore, the quality of GIS data and therefore quality of maps was not assessed quantitatively. Assumption was made that, since the latest and only data from official sources were used, the quality of maps should be good. However, it is necessary to address quality of maps question before any of results could be used for designing diffuse pollution abatement programs.

The model preparation was simplified in order to reduce the preparation time. For instance, land use was used from one year only. Hence, crop rotation was excluded. Although the exclusion of crop rotation increased uncertainty in model results, it was justified due to the fact that the study was intended as a demonstration example rather than a comprehensive abatement solution to the diffuse pollution problem in the Graisupis river catchment. Another simplification was model calibration only for water flows and nitrate loads (excluding calibration for nitrate concentration and phosphorus loads and concentration). This was justified by the intention to use the model only for the estimation of nitrate loading to surface water bodies. Nevertheless, it must be stressed that all these simplifications might be justified in a demonstration example, but to use model results for real decision making one should be very careful. Calibration for phosphates and sediments might be quite important to decision support, since all pollution elements must be addressed simultaneously. The reduction of one parameter should not be counter-weighed by the increase of other pollutants. It is essential that MOSO would address all important water quality elements. Furthermore, it is also important that autocalibration (if applied) would be done simultaneously for all required variables. Although it may not be possible in many cases, it is highly desirable, since autocalibration in steps requires more time and often upsets the calibration results of the previous step.

Definition of CSAs and sensitive areas is essential for their identification. However, there is no clear guidance how to define them. CSAs can be identified based on the impact on water quality, tolerated pollutant loading from the land, pollutant loading deviation from average values, etc. Therefore, before starting CSAs or sensitive areas identification, an agreement between stakeholders should be reached what method is the most suitable to address the problems in a watershed and how it should be applied.

It is also important to address the concept of the SWAT model. SWAT uses HRUs as a unit for modeling. This simplification is useful, if large watersheds are addressed for diffuse pollution modeling. It allows a short run time for the model. However, it becomes a shortcoming of the model, if a precise placement of BEPs is in question. Moreover, the lack of overland routing simulation (SWAT is treating all loading originating anywhere in the HRU the same) creates problems for representation of some BEPs. With some BEPs there are no problems. For instance, conversion to grasslands, afforestation, application of cover crops and residue management could be implemented easily. Yet, it still might be desirable to locate an exact field for the BEP, which is not easy in the HRU concept. Furthermore, the representation of some BEPs, which are aimed at pollution reduction from overland routing, is quite problematic. Although it is possible to represent filter strips with the SWAT model, but to know the exact place where it should be applied or the exact area needed for the installation of filter strips with the current SWAT model is hardly possible.

Costs information is vital for MOSO. Thus, a good economist is no less important than a good watershed modeler. In this study cost information was obtained from the LEPA's Nemunas river district management plan (2010). However, cost information did not include opportunity and other costs. Only installation and maintenance costs were considered. The benefits of each BEP were not included in the study as well. A comprehensive cost-benefit analysis would be desirable for MOSO applications. This information would create more trust in final results.

Despite all the mentioned problems, this study provides some guidance in solving diffuse pollution problems. Until this day there were no suitable proposals how to solve the question of BEPs spatial distribution in Lithuania. This study provides possible answers to that and other questions. For instance, such methods could be used to provide basis to setting restrictions in sensitive areas for agricultural activities or directing funds and efforts towards certain critical areas or designing cost-effective watershed management programs.

However, it should be mentioned that social aspects of diffuse pollution abatement were not touched in this work. Optimization results might provide excellent answers from theoretical perspective, yet they could make no sense within existing social and cultural framework. Therefore the analysis of social factors (such as expectations of stakeholders' group, existing financial initiatives, awareness of environmental problems within stakeholders, desicion making practices, etc) are at least as important as the analysis of environmental factors for solving diffuse pollution problems. It should be understood that methods proposed in this work are not supposed to be used as the single best sollution. Their intention should be more understood as a basis or a starting point for stakeholders' negotiations.

Finally, it is important to discuss methods' relation to geography, since this study is done within the framework of GIS Master's program. This study applies geographical methods for its objectives. The integration of hydrology, which belongs to the branch of physical geography, and environmental management, which comes from the branch of environmental geography, provides powerful composition to deal with diffuse water pollution problems. Methods used in this study applies GIS as a tool for the data and model preparation as well as analysis and presentation of results. Thus this work can be fully encompassed within boundaries of geographical science.

# 8 Conclusions

The problem of diffuse water pollution is the major factor responsible for the deterioration of water quality in Lithuania. However, on the institutional level, there is still no clear answer, how to deal with it. This work has suggested and demonstrated some of the methods, which might be suitable for decision support in diffuse pollution abatement. It is a key moment for these suggestions, since, according to the Water Framework Directive, the implementation of abatement measures should start in 2012 at the latest. Random implementations as well as no implementation of the abatement measures are not good options, since both of them would cause a waste of money and possible fines from the European Union for not reaching the goals of the Water Framework Directive.

Methods suggested by this study included the identification of critical source and sensitive areas and the application of multi-objective spatial optimization in decision support of diffuse pollution abatement. The demonstration of these methods was successful. The identification of critical source areas located key areas responsible for nitrate problems in the Graisupis river catchment. 12.4% of catchment was identified as risk areas. The identification of sensitive areas assigned medium or low sensitivity to most the catchment. This corresponded to soil characteristics in the catchment. However, one important aspect became clear during this study. It is the importance of agreement with stakeholders on the definition of critical source areas and sensitive areas. This is a key step in identifying of critical source and sensitive areas, which should not be ignored.

The application of multi-objective spatial optimization was also successfully done. It increased the cost-effectiveness of pollution abatement 24 times (or up to 50 times when the most cost-effective solution was selected from Pareto optimum) if compared with the random application of best environmental practices. The application of multi-objective spatial optimization also provided optimal placements of best environmental practices, which are important for watershed managers. Genetic algorithms, which also were recommended by similar studies, worked well for multi-objective spatial optimization of this study. Generally, 200 generations were necessary to reach close to optimum solutions. However, 1000 generations allowed to get slight improvements of the results. Pareto optimum, which provided the relationship between economic and environmental objectives, was one of the important results of this study.

Lastly, it is also important to mention that the SWAT model, despite of a few shortcomings, was suitable for the aim of the study. Moreover, available scientific literature provided a lot of ideas and guidelines, how this study could be conducted. Furthermore, it was possible to calibrate and validate the model successfully with the available data and estimated missing parameters. At this time the model was prepared only for the Graisupis river catchment. However, it may be applied for any other place in Lithuania.

# Reference list

Arabi, M., Frankenberger, R. J., Engel, A. B., Arnold, G. J., 2007. Representation of agricultural conservation practices with SWAT. Hydrological processes 22, pp. 3045-3055.

Arabi, M., Govindaraju, S. R., Hantush, M., 2006. Cost-effective allocation of watershed management practices using a genetic algorithm. Water Resources Research 42(W10429), pp. 1-14.

Batjes NH 2008. ISRIC-WISE Harmonized Global Soil Profile Dataset (Ver. 3.1). Report 2008/02, ISRIC – World Soil Information, Wageningen (with dataset). Available at: http://www.isric.org/isric/Webdocs/ISRIC_Report_2008_02.pdf Retrieved: 2010.02.20

Diebel, M. W., Maxted, T. J., Novak, J. P., Vander Zander, J. M., 2008. Landscape planning for agricultural nonpoint source pollution reduction I: A geographical allocation framework. Environmental Management 42, 789-802.

Eawag, 2008. Eawag: Department System Analysis, Integrated Assessment and Modelling: SWAT-CUP. Available at: http://www.eawag.ch/organisation/abteilungen/siam/software/swat/index_EN Retrieved: 2010.04.13.

EDC, 2010. Newcastle University, Engineering Design Centre. The basic algorithm for a GA. Available at: http://www.edc.ncl.ac.uk/highlight/rhjanuary2007g01.php/ Retrieved: 2010.12.30.

Gassman, W. P., Reyes, R. M., Green, H. C., Arnold, G. J., 2007. The Soil and Water Assessment Tool: historical development, applications and future research needs. Transactions of the ASABE 50(4), pp. 1211-1250.

Ghebremichael, L. T., Veith, T. L., Watzin, M. C., 2010. Determination of critical source areas for phosphorus loss: lake Champlain basin, Vermont. Transactions of the ASABE 53(5), pp. 1595-1605.

Gitau, W. M., Veith, L. T., Gburek, J. W., 2004. Farm-level optimization of BMP placement for cost-effective pollution abatement. Transactions of the ASAE. American Society of Agricultural Engineers 47(6), pp. 1923-1931.

Gijsman, J. A., Thornton, K. Ph., Hoogenboom, G., 2007. Using the WISE database to parametrise soil inputs for crop simulation models. Computers and Electronics in Agriculture 56, pp. 85-100.

Huang, J., Hong, H., 2010. Comparative study of two models to simulate diffuse nitrogen and phosphorus pollution in a medium-sized watershed, southeast China. Estuarine, Coastal and Shelf Science 86, pp. 387-394.

Jha, M., Rabotyagov, S., Gassman, W. Ph., 2009. Optimal Placement of Conservation Practices Using Genetic Algorithm with SWAT. Center for Agricultural and Rural Development, Iowa State

University. Working Paper 09-WP 496.

LEPA, 2010. Nemunas river basin district managment plan. Available at: http://vanduo.gamta.lt/cms/index?rubricId=24f1ad10-5558-4a0f-9ac3-4f59720cbcbb Retrieved: 2010.05.30.

Maringanti, Ch., Chaubey, I., Popp, J., 2009. Development of a multiobjective optimization tool for the selection and placement of best management practices for nonpoint source pollution control. Water Resources Research 45(W06406), pp. 1-15.

Mass, R. P., Smolen, M, D., Dressing, S. A., 1985. Selecting critical areas for nonpoint source pollution control. Journal of Soil and Water Conservation 40(1), 68-71.

Meyer, Ch. B., Lescot, M. J., Laplana, R., 2009. Comparison of Two Spatial Optimization Techniques: A framework to solve Multiobjective Land Use Distribution Problems. Environmental Management 43, pp. 264-281.

Minister of Environment Republic of Lithuania, 2010. "Method description for assessment of surface water bodies". /-
-/Lietuvos Respublikos aplinkos ministras (ruošiamos taisyklės). "Paviršinių vandens telkinių būklės vertinimo tvarkos aprašas".

Minister of Environment Republic of Lithuania, 2005. "Approval of description methods for typology of surface water bodies, for water quality indicators required for establishment of reference sites and for criteria on artificial, heavily modified and water bodies at risk." 23/05/2005. /-
-/Lietuvos Respublikos Aplinkos ministro įsakymas (D1-256). Dėl paviršinių vandens telkinių tipų aprašo, paviršinių vandens telkinių kokybės elementų etaloninių sąlygų rodiklių aprašo ir kriterijų dirbtiniams, labai pakeistiems ir rizikos vandens telkiniams išskirti aprašo patvirtinimo. 2005m. gegužės 23 d.

Moriasi, N. D., Arnold, G. J., Van Liew, W. M., Bingner, L. R., Harmel, D. R., Vieth, L. T., 2007. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. American Society of Agricultural and Biological Engineers Vol. 50(3):885-900.

Neitsch, L. S., Arnold, G. J., Kiniry, R. J., Williams, R. J., 2005. Soil and water assessment tool. Theoretical documentation. Version 2005. Grassland, Soil and Water Research Laboratory, Agricultural Reseach Service & Backland Reseach Center, Texas Agricultural Experiment Station. Available at: http://www.brc.tamus.edu/swat/doc.html Retrieved: 2009.02.04.

Neitsch, L. S., Arnold, G. J., Kiniry, R. J., Srinivasan, R., Williams, R. J., 2004. Soil and water assessment tool. Input/Output file documentation. Version 2005. Grassland, Soil and Water Research Laboratory, Agricultural Research Service & Backland Reseach Center, Texas Agricultural Experiment Station. Available at: http://www.brc.tamus.edu/swat/doc.html Retrieved:

2010.02.20.

Noll, R. M., Magee, A. E., 2009. Quantification of phosphorus sources to a small watershed: A case study of Graywood Gully, Conesus Lake, NY. Journal of Great Lakes Research 35, pp. 50-55.
OMAFRA, 2000. Universal Soil Loss Equation (USLE). Website of Ontarion Ministry of Agricultural Food & Rural Affairs. Available at:
http://www.omafra.gov.on.ca/english/engineer/facts/00-001.htm Retrieved on: 2010.05.30.

Ou, Y., Wang, X., 2008. Identification of critical source areas for non-point source pollution in Miyun reservior watershed near Beijing, China. Water Science & Technology 58(11), 2235-2241.

Ouyang, W., Hao, H. F., Wang, L. X., Regional non point source organic pollution modeling and critical area identification for watershed best environment management. Water Air Soil Pollution 187, 251-261.

Perez-Pedini, C., Limbrunner, F. J., Vogel, M. R., 2005. Optimal Location of Infiltration-Based Best Management Practices for Storm Water Management. Journal of Water Resources Planning and Management. 131(6), pp. 441-448.

Plunge, S., 2009. Risks versus Costs: A New Approach for Assessment of Diffuse Water Pollution Abatement. Master of Science Thesis. Chalmers University of Technology, Goteborg, Sweden. Available at:
http://www.chalmers.se/ee/SV/forskning/forskargrupper/miljosystemanalys/publikationer/pdf-filer/2009_1/downloadFile/attachedFile_9_f0/2009-12.pdf?nocache=1266417103.12

Qui, Z., 2009. Assessing Critical Source Areas in Watersheds for Conservation Buffer Planning and Riparian Restoration. Environmental Management 44, pp. 968-980.

Rao, S. N., Easton, M. Z., Schneiderman, M. E., Zion, S. M., Lee, R. D., Steenhuis, S. T., 2009. Modeling watershed-scale effectiveness of agricultural best management practices to reduce phosphorus loading. Journal of Environmental Management 90, pp. 1385-1395.

Sivertun, A., Prange, L., 2003. Non-point source critical area analysis in the Gisselo watershed using GIS. Environmental Modelling & Software 18, 887-898.

Smitiene, A., 2007. Pattern of nitrogen leaching withing streams catchment. Water Management Institute of Lithuanian University of Agriculture./-
-/Šmitienė, A., 2007. Azoto išplovimo vyksmo upelių baseinuose dėsningumai. Dartako disertacija. Lietuvos žemės ūkio universiteto, Vandens ūkio institutas.

Srinivasan, S. M., Gerard-Marchant, P., Veith, L. T., Gburek, J. W., Steenhuis, S. T., 2005. Watershed scale modeling of critical source areas of runoff generation and phosphorus transport.

Journal of the American Water Resources Association 41(2), 361-375.

Strauss, P., Leone, A., Ripa, N. M., Turpin, N., Lescot, M. J., Laplana, R., 2007. Using critical source areas for targeting cost-effective best management practices to mitigate phosphorus and sediment transfer at the watershed scale. Soil Use and Management 23 (Suppl. 1), pp. 144-153.

Trevisan, D., Dorioz, M. J., Poulenard, J., Quetin, Ph., Combaret, P. C., Merot, Ph., 2010. Mapping of critical source areas for diffuse fecal bacterial pollution in extensively grazed watersheds. Water Research 44, pp. 3847-3860.

Tripathi, P. M., Panda, K. R., Raghuwanshi, S. N., 2003. Identification and Prioritisation of Critical Sub-watersheds for Soil Conservation Management using the SWAT Model. Biosystems Engineering 85(3), 365-379.

USDA 1986. Urban Hydrology for Small Watersheds' published by the Engineering Division of the Natural Resource Conservation Service, United States Department of Agriculture, Technical Release–55. Available at: http://www.cset.sp.utoledo.edu/~nkissoff/pdf/CIVE-3520/Modified-tr55.pdf
Retrieved: 2010.02.20

Veith, L. T., Wofle, L. M., Heatwole, D. C., 2003. Optimization procedure for cost effective BMP placement at a watershed scale. Journal of the American Water Resources Association, December, pp. 1331-1343.

Veith, L. T., Wofle, L. M., Heatwole, D. C., 2004. Cost-Effective BMP Placement: Optimization versus Targeting. American Society of Agricultural Engineers 47(5), pp. 1585-1593.

Watershedss, 2010. Critical Areas in Agricultural Nonpoint Source Pollution Control Projects. Available at: http://www.water.ncsu.edu/watershedss/dss/estuary/shellfish/five.html Retrieved on: 2010.05.30.

White, J. M., Storm, E. D., Busteed, R. P., Stoodley, H. S., Phillips, J. Sh., 2009. Evaluation Nonpoint Source Critical Source Area Contributions at the Watershed Scale. Journal of Environmental Quality 38, 1654-1663.

WMI, 2009. The Analysis of Land Use, Chemical Composition of Water and Precipitation in Typical Agro-ecosystems of Middle and West Lithuania. General report on agro-ecosystem research. WaterManagement Institute of Lithuanian University of Agriculture /-
-/VŪI, 2009. Žemėnaudos, vandens ir kritulių cheminės sudėties ir savybių tyrimai tipiškose vidurio ir vakarų Lietuvos agroekosistemose. Bendroji agroekosistemų tyrimų ataskaita. Lietuvos žemėsūkio universiteto Vandens ūkio institutas. Prieinamas:
http://gamta.lt/files/Agroekosistemos_tyrimai_2008m.pdf

WMI, 2007. The Analysis of Land Use, Chemical Composition of Water and Precipitation in Typical Agro-ecosystems of Middle and West Lithuania. General report on agro-ecosystem research. Water Management Institute of Lithuanian University of Agriculture /-
-/VŪI, 2007. Žemėnaudos, vandens ir kritulių cheminės sudėties ir savybių tyrimai tipiškose vidurio ir vakarų Lietuvos agroekosistemose. Bendroji agroekosistemų tyrimų ataskaita. Lietuvos žemės ūkio universiteto Vandens ūkio institutas. Prieinamas:
http://gamta.lt/files/2006m1240910669227.pdf

# Appendix A.

| National soil classification system (soil unit symbol) | FOA soil classification system (soil unit symbol) | FOA soil classification system |
|---|---|---|
| ADb | FLe | Eutric Fluvisols |
| GLk | GLk | Calcic Gleysols |
| GLv | GLm | Mollic Gleysols |
| IDg | LVg | Gleyic Luvisols |
| IDk | LVk | Calcic Luvisols |
| IDp | LVh | Haplic Luvisols |
| PLb | PLe | Eutric Planosols |
| RDg | CMg | Gleyic Cambisols |
| RDk | CMc | Calcaric Cambisols |
| SDg | ARg | Gleyic Arenosols |
| SDp | ARh | Haplic Arenosols |
| SDk | ARc | Calcaric Arenosols |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | FLe | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | A | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.62 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.15 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.58 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 448.26 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 8 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 12 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 80 | HWSD |

| | | | |
|---|---|---|---|
| ROCK (% wt.) | Rock fragment content (% total weight) | 18 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.13 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.59 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.15 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.24 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 390.69 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 10 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 11 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 79 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 10 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.13 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | GLk | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.42 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.125 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 1.3 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 28.1 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 19 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 40 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 41 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |

| | | | |
|---|---|---|---|
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.15 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.42 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.125 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.44 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 32.9 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 19 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 37 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 44 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 5 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.18 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | GLm | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.39 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.125 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 1.65 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 21.19 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 22 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 39 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 39 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |

| | | | |
|---|---|---|---|
| USLE_K | USLE equation soil erodibility (K) factor | 0.14 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.35 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.125 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.69 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 15.44 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 28 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 35 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 37 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 3 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.16 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | LVg | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.39 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.83 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 33.01 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 24 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 29 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 47 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 5 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.16 | Estimated |

| | | | |
|---|---|---|---|
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.33 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.28 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 16.75 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 34 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 27 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 39 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 6 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.16 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | LVk | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | C | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.41 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.51 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 50.74 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 23 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 24 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 53 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 5 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.15 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Layer Parameters / Layer 2 | | | |
|---|---|---|---|
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.36 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.3 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 30.74 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 30 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 23 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 47 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 10 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.15 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.2 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | LVh | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.4 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.15 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.74 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 23.84 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 22 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 37 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 41 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.17 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |

| | | | |
|---|---|---|---|
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.35 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.15 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.36 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 15.19 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 29 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 34 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 37 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 3 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.17 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | PLe | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.41 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.075 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 1.06 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 27.85 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 20 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 38 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 42 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.18 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |

| | | | |
|---|---|---|---|
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.3 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.075 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.43 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 10.68 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 36 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 31 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 33 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.16 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | CMg | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.42 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 1 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 34.76 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 19 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 36 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 45 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.16 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.36 | HWSD |

| | | | |
|---|---|---|---|
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.47 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 18.05 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 39 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 34 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 27 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 18 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.17 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | CMc | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | B | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.39 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.05 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.65 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 19.02 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 21 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 43 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 36 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 6 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.17 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.4 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.38 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm | 0.05 | HWSD |

| | | | |
|---|---|---|---|
| | soil) | | |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.43 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 15.15 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 23 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 43 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 34 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 10 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.18 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.3 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | ARg | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | A | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.7 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.015 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.5 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 822.95 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 7 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 88 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 3 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.09 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.7 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.015 | HWSD |

| | | | |
|---|---|---|---|
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.23 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 873.56 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 6 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 89 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.09 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | ARh | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | A | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.71 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.015 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.4 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 928.11 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 5 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 90 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.09 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.08 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.71 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.015 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.21 | HWSD |

| | | | |
|---|---|---|---|
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 928.11 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 5 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 90 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.08 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

| Soil Component Parameters | Definition | Value | Data source |
|---|---|---|---|
| SNAM | Soil name | ARc | HWSD |
| LAYERS | Number of layers | 2 | HWSD |
| HYDGRP | Soil hydrologic group (A,B, C, or D) | A | HWSD |
| SOL_ZMX (mm) | Maximum rooting depth of soils profile (mm) | 1000 | HWDS |
| ANION_EXCL (fraction) | Fraction of porosity (void space) from which anions are excluded | 0.5 | Default |
| SOL_CRK (m3/m3) | Potential or maximum crack volume of the soil profile expressed as fraction of the total soil profile | 0.5 | Default |
| Soil Layer Parameters / Layer 1 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 300 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.7 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.4 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 873.56 | Estimated |
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 6 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 89 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 3 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.09 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |
| Soil Layer Parameters / Layer 2 | | | |
| SOL_Z (mm) | Depth from soil surface to bottom of layer (mm) | 1000 | HWSD |
| SOL_BD (g/cm3) | Moist bulk density (Mg/m3 or g/cm3) | 1.71 | HWSD |
| SOL_AWC (mm/mm) | Available water capacity of the soil layer (mm H2O/mm soil) | 0.1 | HWSD |
| SOL_CBN (% wt.) | Organic carbon content (% soil weight) | 0.2 | HWSD |
| SOL_K (mm/hr) | Saturated hydraulic conductivity (mm/hr) | 928.11 | Estimated |

| | | | |
|---|---|---|---|
| CLAY (% wt.) | Clay content (% soil weight) | 5 | HWSD |
| SILT (% wt.) | Silt content (% soil weight) | 5 | HWSD |
| SAND (% wt.) | Sand content (% soil weight) | 90 | HWSD |
| ROCK (% wt.) | Rock fragment content (% total weight) | 4 | HWSD |
| SOL_ALB (fraction) | Moist soil albedo | 0.13 | Estimated |
| USLE_K | USLE equation soil erodibility (K) factor | 0.08 | Estimated |
| SOL_EC (dS/m) | Electrical conductivity (dS/m) | 0.1 | HWSD |

# Appendix B.

## The algorithm code of cost-effectiveness optimization

```
Sub cellset(zSheetname, zRow, zCol, zVal)
Worksheets(zSheetname).Cells(zRow, zCol) = zVal
End Sub
Sub Optimization()
'Stating variables and constants
Dim ad(1 To 25) As Double
Dim population(1 To 604, 1 To 100) As Integer
Dim r As Range
Dim x As Integer
Dim crossover_point As Integer
Dim parent1 As Integer
Dim parent2 As Integer
Dim random As Double
cost_ef = 0
cost_ef_fix = 99999999
m = 1
'GA parameters
pop_size = 100
generation_nb = 1000
mutation_rate = 0.001
'Creating initial population with random genome
For ind = 1 To pop_size
   For gene = 1 To 604
      cellset "POP", gene, ind, Int((6 - 1 + 1) * Rnd + 1)
   Next
Next
For generation = 1 To generation_nb
   'Calculating load reduction and costs for each individual
   For ind = 1 To pop_size
      loads = 0
      costs = 0
      For gene = 1 To 604
         allele = Worksheets("POP").Cells(gene, ind)
         loads = loads + Worksheets("LOADS").Cells(gene + 1, allele + 1)
         costs = costs + Worksheets("COSTS").Cells(gene + 1, allele + 1)
      Next
      cellset "STAT", ind, 1, ind
      cellset "STAT", ind, 2, loads
      cellset "STAT", ind, 3, costs
      cost_ef = costs / (22487 - loads)
      'Fixing the best solution obtained during optimization
      If cost_ef_fix > cost_ef Then
         For gene = 1 To 604
            cellset "BEST", gene, 1, Worksheets("POP").Cells(gene, ind)
         Next
         cost_ef_fix = cost_ef
         cellset "BEST", 1, 2, cost_ef_fix
         cellset "BEST", 1, 3, loads
         cellset "BEST", 1, 4, costs
         cellset "BEST", 1, 5, generation
```

```
      End If
   Next
   x = 1
   'Ranking is done and updated automaticly in Excel spreadsheet. It is done in the following steps:
      'First, following formula is used to get average of rank between load reduction and costs //
(RANK(B1,$B$1:$B$100)+RANK(C1,$C$1:$C$100))/2
      'Second, results are ranked once again to abtain overall rank // RANK(D1,$D$1:$D$100)
   '25% of population is selected into breeding pool
   For ind = 1 To 100
      If Worksheets("STAT").Cells(ind, 5) < 26 And x < 26 Then
      ad(x) = Worksheets("STAT").Cells(ind, 1)
      cellset "STAT", x, 7, ad(x)
      x = x + 1
         'Recording improvement in the evolution of cost effectiveness
         If generation > generation_nb - 100 Then
            cellset "SAVE", m, 1, Worksheets("STAT").Cells(ind, 2)
            cellset "SAVE", m, 2, Worksheets("STAT").Cells(ind, 3)
            m = m + 1
         End If
      End If
   Next
   'Creating a new population
   For ind = 1 To pop_size
      'Selection of parents
      x = Int((25 - 1 + 1) * Rnd + 1)
      parent1 = ad(x)
      x = Int((25 - 1 + 1) * Rnd + 1)
      parent2 = ad(x)
      crossover_point = Int((604 - 1 + 1) * Rnd + 1)
      'Crossover
      For gene = 1 To crossover_point
         random = Rnd
         If random <= mutation_rate Then
            population(gene, ind) = Int((6 - 1 + 1) * Rnd + 1)
         Else
            population(gene, ind) = Worksheets("POP").Cells(gene, parent1)
         End If
      Next
      'Mutation of genes
      For gene = crossover_point + 1 To 604
         random = Rnd
         If random <= mutation_rate Then
            population(gene, ind) = Int((6 - 1 + 1) * Rnd + 1)
         Else
            population(gene, ind) = Worksheets("POP").Cells(gene, parent2)
         End If
      Next
   Next
   'Fixing new population
   For ind = 1 To pop_size
      For gene = 1 To 604
         cellset "POP", gene, ind, population(gene, ind)
      Next
   Next
   'Fixing the load reduction and cost information of each generation
   cellset "STAT", generation, 11, Worksheets("STAT").Cells(1, 6)
   cellset "STAT", generation, 12, Worksheets("STAT").Cells(2, 6)
   cellset "STAT", 3, 6, Worksheets("STAT").Cells(generation, 12) / (22487 - Worksheets("STAT").Cells(generation,
```

```
11))
   cellset "STAT", generation, 13, Worksheets("STAT").Cells(generation, 12) / (22487 -
Worksheets("STAT").Cells(generation, 11))
   cellset "STAT", 4, 6, generation
Next
End Sub
```

## The algorithm code of Pareto optimum optimization

```
Sub cellset(zSheetname, zRow, zCol, zVal)
Worksheets(zSheetname).Cells(zRow, zCol) = zVal
End Sub
Sub Optimization()
'Stating variables and constants
Dim ad(1 To 25) As Double
Dim population(1 To 604, 1 To 100) As Integer
Dim r As Range
Dim x As Integer
Dim crossover_point As Integer
Dim parent1 As Integer
Dim parent2 As Integer
Dim random As Double
cost_ef = 0
cost_ef_fix = 99999999
'GA parameters
pop_size = 100
generation_nb = 3000
mutation_rate = 0.001
'All Pareto optimization should be rerun 3 times
For loop_basis = 1 To 3
m = 1
'Creating initial population with random genome
   For ind = 1 To pop_size
      For gene = 1 To 604
         cellset "POP", gene, ind, Int((6 - 1 + 1) * Rnd + 1)
      Next
   Next
   'Calculating load reduction and costs for each individual
   For generation = 1 To generation_nb
      For ind = 1 To pop_size
         loads = 0
         costs = 0
         For gene = 1 To 604
            allele = Worksheets("POP").Cells(gene, ind)
            loads = loads + Worksheets("LOADS").Cells(gene + 1, allele + 1)
            costs = costs + Worksheets("COSTS").Cells(gene + 1, allele + 1)
         Next
         cellset "STAT", ind, 1, ind
         cellset "STAT", ind, 2, loads
         cellset "STAT", ind, 3, costs
         cost_ef = costs / (22487 - loads)
         'Fixing the best solution obtained during optimization
         If cost_ef_fix > cost_ef Then
            For gene = 1 To 604
               cellset "BEST", gene, 1, Worksheets("POP").Cells(gene, ind)
            Next
            cost_ef_fix = cost_ef
```

```
            cellset "BEST", 1, 2, cost_ef_fix
            cellset "BEST", 1, 3, loads
            cellset "BEST", 1, 4, costs
            cellset "BEST", 1, 5, generation
        End If
    Next
    x = 1
    'Ranking is done and updated automaticly in Excel spreadsheet. It is done in the following steps:
        'First, following formula is used to get average of rank between load reduction and costs //
(RANK(B1,$B$1:$B$100)+RANK(C1,$C$1:$C$100))/2
        'Second, results are ranked once again to abtain overall rank // RANK(D1,$D$1:$D$100)
    '25% of population is selected into breeding pool
    '///////////////////////////
    'For the first thousand generations optimization is done as in the first optimization
    If generation <= generation_nb - 2000 Then
        For ind = 1 To 100
            If Worksheets("STAT").Cells(ind, 5) < 26 And x < 26 Then
                ad(x) = Worksheets("STAT").Cells(ind, 1)
                cellset "STAT", x, 7, ad(x)
                x = x + 1
            End If
        Next
    End If
    '/////////////////////////////////
    'For the second thousand generations optimization is examining the effect of load reduction decrease on costs
    If generation > generation_nb - 2000 And generation <= generation_nb - 1000 Then
        'The population of 2001 generation is saved seperately
        If generation = generation_nb - 1999 Then
            For ind = 1 To pop_size
                For gene = 1 To 604
                    cellset "POP_SAVE", gene, ind, Worksheets("POP").Cells(gene, ind)
                Next
            Next
        End If
        For ind = 1 To 100
            If Worksheets("STAT").Cells(ind, 2) > Worksheets("STAT").Cells(1, 6) Then
                If Worksheets("STAT").Cells(ind, 5) < 26 And x < 26 Then
                    ad(x) = Worksheets("STAT").Cells(ind, 1)
                    cellset "STAT", x, 7, ad(x)
                    x = x + 1
                    cellset "SAVE", m, loop_basis, Worksheets("STAT").Cells(ind, 2)
                    cellset "SAVE", m, loop_basis + 3, Worksheets("STAT").Cells(ind, 3)
                    m = m + 1
                End If
            End If
        Next
        If x < 25 Then
            For ind = 1 To 100
                If Worksheets("STAT").Cells(ind, 5) < 51 And x < 26 And Worksheets("STAT").Cells(ind, 2) >
Worksheets("STAT").Cells(1, 6) Then
                    ad(x) = Worksheets("STAT").Cells(ind, 1)
                    cellset "STAT", x, 7, ad(x)
                    x = x + 1
                End If
            Next
        End If
    End If
    '////////////////////////
```

```
'For the third thousand generations optimization is examining the effect of load increase on costs
If generation > generation_nb - 1000 Then
   If generation = generation_nb - 999 Then
      For ind = 1 To pop_size
         For gene = 1 To 604
            cellset "POP", gene, ind, Worksheets("POP_SAVE").Cells(gene, ind)
         Next
      Next
      For ind = 1 To pop_size
         loads = 0
         costs = 0
         For gene = 1 To 604
            allele = Worksheets("POP").Cells(gene, ind)
            loads = loads + Worksheets("LOADS").Cells(gene + 1, allele + 1)
            costs = costs + Worksheets("COSTS").Cells(gene + 1, allele + 1)
         Next
         cellset "STAT", ind, 1, ind
         cellset "STAT", ind, 2, loads
         cellset "STAT", ind, 3, costs
      Next
   End If
   For ind = 1 To 100
      If Worksheets("STAT").Cells(ind, 2) < Worksheets("STAT").Cells(1, 6) Then
         If Worksheets("STAT").Cells(ind, 5) < 26 And x < 26 Then
            ad(x) = Worksheets("STAT").Cells(ind, 1)
            cellset "STAT", x, 7, ad(x)
            x = x + 1
            cellset "SAVE", m, loop_basis, Worksheets("STAT").Cells(ind, 2)
            cellset "SAVE", m, loop_basis + 3, Worksheets("STAT").Cells(ind, 3)
            m = m + 1
         End If
      End If
   Next
   If x < 25 Then
      For ind = 1 To 100
         If Worksheets("STAT").Cells(ind, 5) < 51 And x < 26 And Worksheets("STAT").Cells(ind, 2) <
Worksheets("STAT").Cells(1, 6) Then
            ad(x) = Worksheets("STAT").Cells(ind, 1)
            cellset "STAT", x, 7, ad(x)
            x = x + 1
         End If
      Next
   End If
End If
'///////////////////////////////
'Creating a new population
For ind = 1 To pop_size
   'Selection of parents
   x = Int((25 - 1 + 1) * Rnd + 1)
   parent1 = ad(x)
   x = Int((25 - 1 + 1) * Rnd + 1)
   parent2 = ad(x)
   crossover_point = Int((604 - 1 + 1) * Rnd + 1)
   'Crossover
   For gene = 1 To crossover_point
      random = Rnd
      If random <= mutation_rate Then
         population(gene, ind) = Int((6 - 1 + 1) * Rnd + 1)
```

```
            Else
               population(gene, ind) = Worksheets("POP").Cells(gene, parent1)
            End If
         Next
         'Mutation of genes
         For gene = crossover_point + 1 To 604
            random = Rnd
            If random <= mutation_rate Then
               population(gene, ind) = Int((6 - 1 + 1) * Rnd + 1)
            Else
               population(gene, ind) = Worksheets("POP").Cells(gene, parent2)
            End If
         Next
      Next
      'Fixing new population
      For ind = 1 To pop_size
         For gene = 1 To 604
            cellset "POP", gene, ind, population(gene, ind)
         Next
      Next
      'Fixing the load reduction and cost information of each generation
      cellset "STAT", generation, 11, Worksheets("STAT").Cells(1, 6) 'average loads
      cellset "STAT", generation, 12, Worksheets("STAT").Cells(2, 6) 'average cost
      cellset "STAT", 3, 6, Worksheets("STAT").Cells(generation, 12) / (22487 - Worksheets("STAT").Cells(generation,
11))
      cellset "STAT", generation, 13, Worksheets("STAT").Cells(generation, 12) / (22487 -
Worksheets("STAT").Cells(generation, 11))
      cellset "STAT", 4, 6, generation
   Next
Next
End Sub
```

Series from Lund University's Geographical Department

**Master Thesis in Geographical Information Science (LUMA-GIS)**

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland. (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research. (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing. (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia. (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010)
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010)
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010)
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011)
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011)