GIMA

Geographical Information Management and Applications

DISTANCE: A DECISIVE GEOSPATIAL FACTOR FOR IDENTIFYING OPTIMAL SUMMER VACATION LOCATIONS. CASE STUDY OF PAROS

Master Thesis Research

Eleni Boboti

Universiteit Utrecht

TU Delft

ITC

Wageningen UR

DISTANCE: A DECISIVE GEOSPATIAL FACTOR FOR IDENTIFYING OPTIMAL SUMMER VACATION LOCATIONS.
CASE STUDY OF PAROS

Master Thesis Research
September 2011

Author: Eleni Boboti

Supervisor: Ron van Lammeren; Wageningen University

Professor: Arnold Bregt; Wageningen University

Reviewer: Fred Toppen; Utrecht University

Abstract

Many people intend to go for summer holidays, and also everyone keeps on planning these holidays almost every year whether it is a single person, a couple or a family. But how do these people decide on which is going to be the most suitable destination for them? What are they looking for when concluding of where to go? How does distance to certain locations can influence their decisions and how does it contribute on these decisions? This is what this thesis is going to be dealing with, giving answers to these questions.

There are a lot of websites where help is provided to tourists who do not have a specific idea in their mind on which can be a pleasant and satisfying destination for their precious, and a lot of times limited, holidays. Even though the existing sites provide some solutions according to what tourists are interested in, they are all location based. One can only check if a certain activity or place is present at a certain area or not. However, this is not the only important aspect. Tourists are also willing to 'travel' from their hotel up to a certain distance in order to reach a certain place of interest. For this reason, the spatial component of distance to specific locations should also be considered as an essential factor when choosing a holiday destination. In our case, we will describe the different aspects of distance and the ways that it can be calculated in order to give valuable results in selecting a vacation destination.

Through this research we get to know what is really most important for a summer tourist while he is on vacation. We will gather information about which places are mostly wanted and based on that, we will develop an application which will provide individual users with the most suitable vacation locations according to how close these users want to be to certain locations of interest once they get to their potential destination. This way we will start building an application part which will later on be a very useful tool, and more advanced from the already existing ones, which will be a more realistic help for tourists. A comparison of two different approaches, simple buffer and network based distances, will also be present in order to see the differences between two different ways of tackling the problem. And of course at the end of this paper there will be a discussion about the results as well as some recommendations.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

APPENDICES

1. INTRODUCTION

Tourism has always played a very important role in some countries' economies. People from around the world work the whole year to be able to have some holidays when they can spend some time for themselves to relax together with their beloved people, their family or even alone. Many countries have developed strategies and plans in order to attract as many tourists as possible. Especially those which lack of big industrial activities, they usually focus on developing the touristic sector.

Planning is necessary for tourism development for many reasons. First of all, to be able to set some goals which later on need to be reached. These objectives include concrete activities which have to be executed in order to bring some results. For tourism the planning goals include in general visitor satisfaction, improving the economy of a nation and area integration to name a few (Gunn, 2002).

Since tourism is strongly related to space, GIS could be of big importance to generate useful information regarding all the planning. In general, the geographers started to work on tourism issues from about the '30s but a combination of geography and tourism issues were written in literature only in the early '60s (Pearce, 1979).

Greece is an example of a country that financially counts a lot on touristic attractions. There are many places in the country that one may be interested in visiting, from ancient settlements, to monastery states, to forest areas to cities with a lot of cultural background to numerous beautiful islands for summer vacation and not only. A lot of people have already visited Greece or they plan on visiting the country. A proof of this interest can be seen through many touristic guides and vacation web sites where Greece is in most of the cases present as an option.

Even though tourism is always connected to leisure and a nice time for the tourists, there is a lot of work to be done on the background. The providers of tourist accommodation/ activities etc. need to be able to propose suitable destinations for the candidate tourists. There are a lot of factors that one should have in mind in order to make some useful conclusions regarding tourism destinations. These would be a blend of internal and external factors as stated by Tzu-Kuang Hsu et al. The internal factors include leisure, adventure, relaxation etc. whereas external factors consist of recreational activities, cultural attractions and so on. (Tzu-Kuang Hsu et al., 2009) But apart from these factors, spatial reasons are also influencing one's decision. Nowadays, with the developments of spatial related tools, it is possible to examine some spatially related reasons why people might choose a certain location for their vacation.

In general, there will be an investigation on which would be the ideal destination for a certain group of people to spend their vacation at. Already on first thought, this can be a result of existing facilities or places of interest, as well as landscape characteristics. In addition, spatial issues can contribute in forming the best area for vacation. The proximity to certain locations can be an example. Such characteristics form part of the behaviour of a tourist which is multidimensional (Boniface, Cooper, 2009). In this research we will focus on building a part of a distance driver application which will help tourists decide on which area is the most suitable for them to go on vacation based on distances to certain locations of interest. Gathering all the needed information will provide us with useful results which we can, later on, use to form more crisp criteria for evaluating the suitability of a certain area.

In our case, we will focus on finding optimal tourism locations using a multi-criteria analysis method. According to H. Zhang and G.H. Huang, 'the GIS-based multi-criteria analysis (GIS-MCA) method combines geographical data and value judgments to obtain information for decision making'. This method (MCA) has also been considered to be a method that helps GIS be part of a decision-support tool (Beedasy J., Whyatt D., 1999). Finding out about the use and possibilities of multi-criteria analysis methods will also help us get to know how a GIS method can be useful in calculating rather precise areas which combine a number of customized preferences for mostly vacations. According to Jacek Malczewski (1999), there are a lot of different approaches to estimate such areas but in this case we want to know what the contribution of this specific method would be.

1.1 PROBLEM DEFINITION

Until this moment tourists plan their holidays according to what they hear from other people, being socially driven, and most of the times having in mind a few interesting things that they would like to do while they are on vacation. People usually chose to stay to places where most of the interesting destinations are close by (Lin G. et.al; 2002). But how do they know whether the place that they are going to visit fulfils as many as possible of their expectations? How do they know if they can have a combination of what they need? It happens a lot of times that a tourist thinks: 'I would like to go somewhere that there is a beach within a certain distance from my hotel. Also there has to be a night club within a decent driving distance because I would like to go out at night.' And of course every person has her or his own preferences which in combination will make the perfect holiday destination for them. It is rather obvious that such expectations consist of spatial components.

There are already some applications which support tourists in their decisions. The *Magic Tour* was an authoring system which gave the basis for tourism applications including multimedia and GIS technologies (Camara, Raper 1999). Also on the internet there are already some applications which make the tourists' life easier in selecting a destination. The air company Lufthansa (www.lufthansa.com) for example, on their website already includes a small selection of preferences in order to narrow down and propose the most suitable destination for the passenger (see Figure 1.1). The problem in this application is that there is a limited spatial aspect in it. We would like to give the tourist the opportunity to state that he or she prefers to stay in a certain minimum or maximum distance from certain locations or/ sites. Lufthansa's example is only offering a point selection and therefore it doesn't really help the tourist find out if he is going to have certain locations close to where he stays or not.

Figure1.1: Lufthansa's trip finder
Source: www.lufthansa.com

Alterra, a research institute in Holland, has already made a very interesting try to come up with suggestions for people who plan on visiting Holland, according to some preference inputs to their online application (http://www.daarmoetikzijn.nl/) which is shown in figure 1.2.

Unfortunately, this application even though it is also based on a multi-criteria analysis, it still does not take into consideration any distance related calculations. It is again an example where suitable areas are shown according to whether these areas have what a tourist needs or not. It is more or less the same kind of an application with Lufthansa's with two major differences. The first one is that the result is areas instead of points and the fact that weights are also assigned to the criteria that have been taken into consideration.

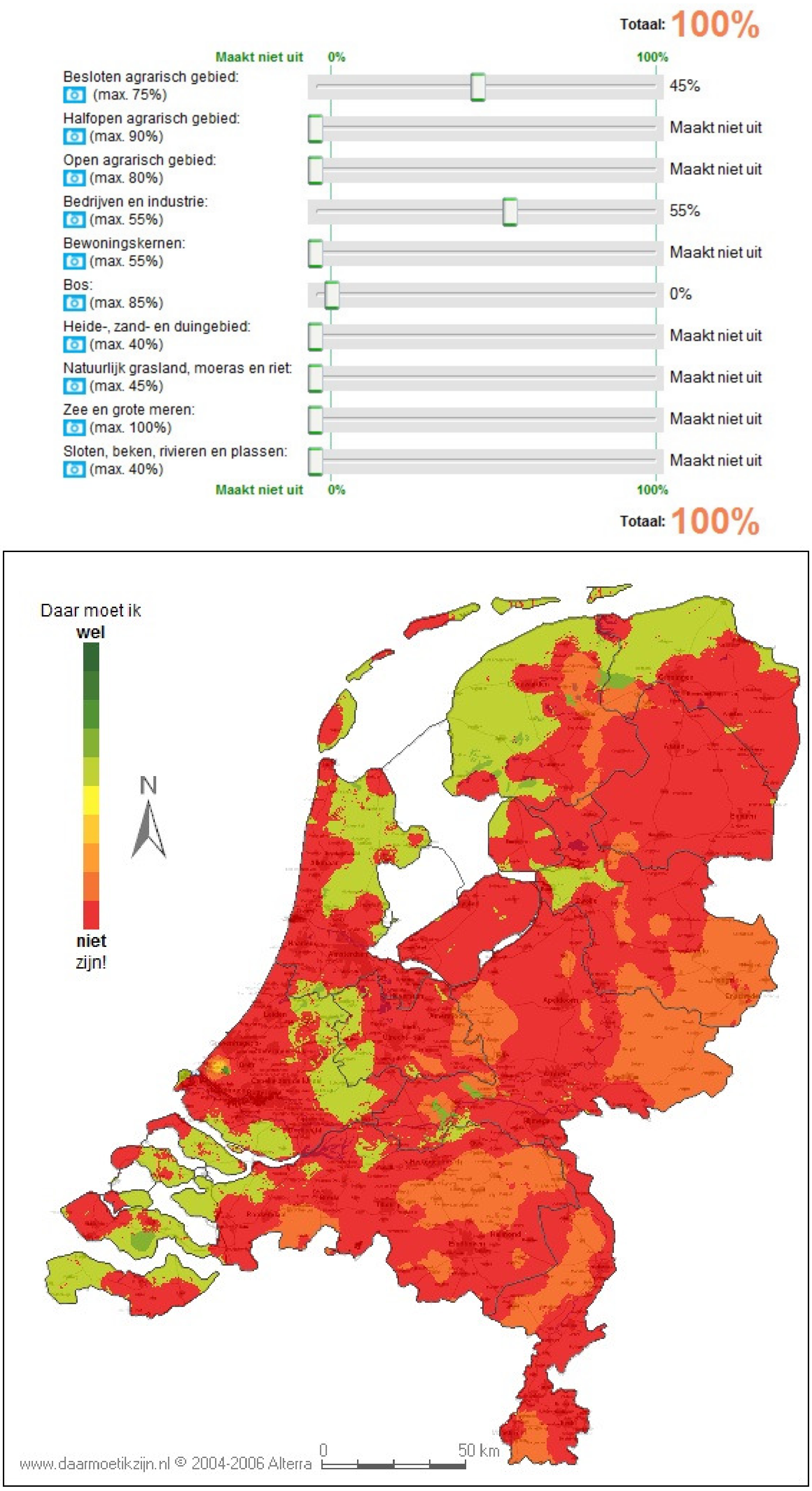

Figure 1.2: An example of area suggestions in Holland.
http://www.daarmoetikzijn.nl/

In general, one could say that he wants to have certain facilities at a certain location. One can rightly claim that this is not something new, that it is present in the already existing applications. For this reason in this research we will also include the spatial preference if distance to certain places of interest. For example a tourist might want to be within a walking distance from a certain beach or a hotel. Also they might be interested in visiting an area where the terrain is not extreme so that they could go for cycling trips with their children. These are a few examples which can form the final suitable areas but it is definitely an issue that needs to be examined in further detail in order to conclude on which factors are most important.

There will be an investigation on which places could be suitable for a group of tourists to be their priority listing in suitable destinations for summer holidays according to their preferences. The difference in our approach is that the spatial criterion of distance is also going to be present in our part of the application and therefore in the result. At the same time it would also be worth it to estimate the places where some people would be least interested in visiting. The important/interesting sites, in general, for a tourist will be narrowed down to a few, and then an application will be built in order to help every individual specify what exactly his expectations are from where he is about to stay as a tourist. This could be anything, from distance to the nearest restaurant, to how steep they want it to be, in case they are cycling, to how much they want to have direct sun. Or course this could include an enormous number of different distance depending on the categories of people.

1.2 RESEARCH OBJECTIVES

Like any other research, it is essential to have an initial overview of what this research is all about what the objectives are and what we will be working on in order to find a solution to the problem specified above.

Our main objective is to estimate optimal summer vacation locations based on physical distances to certain locations. For the purposes of this research, an application will be built and tested, where a single user could specify the importance of the criteria according to his personal preferences. This single user is going to be a person from a certain formed preference group. We will develop a part of a distance based GIS application which will provide solutions about which places of Paros island are most suitable for a tourist to go for summer vacations. In general, this will be achieved by combining distance preferences of a tourist to certain places of interest. He or she will get as a result a highlight of the areas of Paros which are within a specified distance to certain locations.

From the above mentioned objective we can form the research questions of this study which will help us deal with the problem.

Questions:

1. What makes a destination suitable for holidays?
    a. Specify tourist's preferences
    b. How can preferences influence the destination selection?

2. How can distance analysis help in finding tourism destinations?
    a. What are the aspects of distance?
    b. How to implement the spatial factor of distance in calculating optimal destinations?

1.3 THESIS LAYOUT

This thesis is structured in a way that the reader will go through it and little by little get into the point of the research and to the results of it. Even if somebody is not very familiar with the topic, the explanations start from a rather basic stage. The whole document is divided in 8 chapters which will be described in the following paragraphs.

The first two chapters, chapter 1 and 2, include a general description of what the problem is, what we have in mind of doing in order to solve it and which methodology we will make use of to do so. The problem is broken into pieces, the research questions, which are going to be our guides for tackling the problem with a structure. Once we state what our goal is then we will present our overall methodological ways to work on the solution of the problem and that will be in chapter 2.

In chapter three, the important to this thesis factor of distance and some concepts of calculating it are explained. This chapter will be the base to show how the special spatial factor of distance can be used to tackle our problem.

Chapter 4 will deal with forming a certain target group of tourists which will be the sample that is going to provide us with useful information about what tourists want. After concluding about which kind of people will be taken into account we will focus on getting to know from them what is important for tourists and this is going to be shown in chapter 5. We will see that via different ways of collecting information we will come to some conclusion about this matter also. Once we will gather the preferences of tourists, we have to check if out case study area of Paros can actually offer what the tourists ask for. Otherwise there would be no point in starting to analyse which areas of the island are suitable in more detail. This whole issue will be examined in chapter 6.

Chapter 7 will explain the way our specific analysis works and what we will do, in more technical detail, for calculating optimal summer vacation locations in Paros as well as the results of it.

The last chapter will state our overall conclusions about all the issues that were raised in the previous chapters and also a discussion and further recommendations for such a topic.

2. METHODOLOGY

In this chapter, the methodology to reach our final goal is being investigated. As stated in chapter one, the main question is what is the most suitable place for a tourist to go on vacation. In order to do this we would generally have to go through two steps. First to find out which places (facilities, amenities) would be of interest to most of the tourists and then how important each facility is for them (Zhu, X et. al, 2005). This will be a result of finding out the preferences of the tourists and it should give an answer to the question of specifying the combination of the important preferences of the tourists. The second step is to find out in what way the parameter of distance to certain locations of interest. We will show how we will gather information about tourist preferences as well as what the concepts are for developing our part of a general tourist consulting application.

To make things a bit clearer from the beginning, figure 2.1 shows the steps which we will take in order to reach our final results in order to suggest suitable areas for summer vacations to tourists.

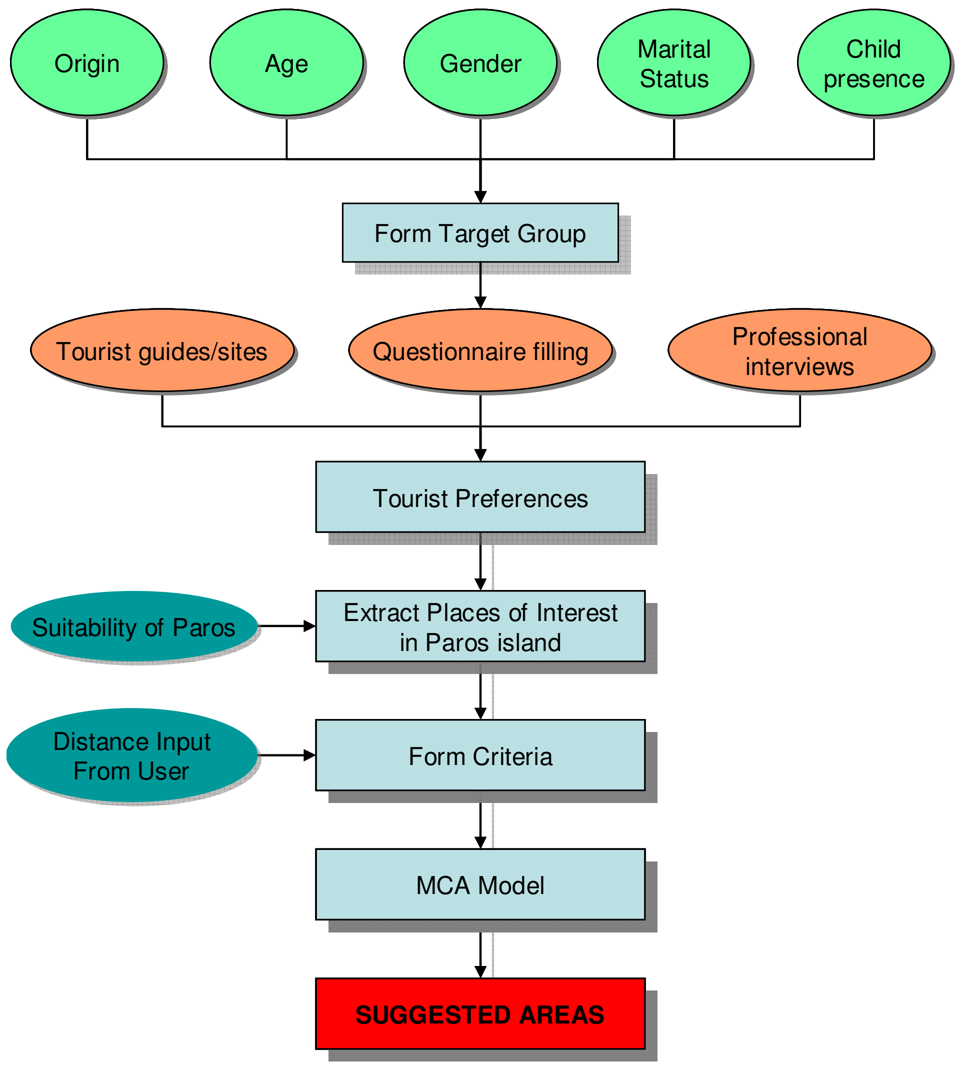

Figure 2.1: The methodology parts.

In general, the rectangles of this diagram show actions whereas the oval shapes show inputs. In the green colour is the input which will be considered in order to form a target group which will later on be asked about their preferences. The orange colour shows the inputs we will have in order to find out the preferences of the tourists. And finally, the dark blue coloured ovals show some additional inputs which are necessary to proceed with calculating the final suggested areas.

2.1 GATHERING AND SPECIFYING TOURIST'S PREFERENCES

As seen in the workflow what we need to start with is setting a tourist target group so that we narrow down to a specific group of people whose opinion and preferences will be taken into consideration. This group will be formed according to certain criteria such as specifying an age group, the origin etc. People who fulfil all the criteria of the target group will then be asked to fill out a certain questionnaire

(Appendix 2). By using the results of this questionnaire in combination with information from tourist guides, websites and professional interviews, such as hotel managers and tourist office personnel, we can then conclude on which places are most important for the tourists. By checking if Paros island can offer these preferred places we can find out what are the specific places of interest and where these are located in the specific case of Paros island. For example we will know that beaches are very important for tourists and therefore we will show where there are beaches in Paros. The same thing will happen with other places of interest as well. After users specify the maximum distance within which they want to be from those places of interest we will have our criteria for our multi-criteria analysis model. Every criterion will be the distance to a certain kind of place. By combining these criteria there will be a calculation of which areas in Paros island fulfil all or some of the criteria, giving the tourist an overview of which areas are best for him to stay at during their summer vacations.

As part of this research, different approaches for solving this problem will be investigated. For example how different it would be if we had a static solution, where all the preferences are preset and no user can influence the result, over a dynamic one which allows users to create a result according to their own personal choices. What would be the difference if our model was predefined and static instead of dynamic were each person can set his own preferences to an application. Also parts within the process of deciding via a multi-criteria method will have to be adjusted according to the needs of the study (Malczewski J., 1999).

Since this thesis is rather focused on a spatial point of view, the most important aspect is to find out the geographical/spatial reasons why people choose certain destinations, which are their preferences and most particularly the distance preferences to certain locations. There are many sources for collecting this kind of information such as tourist websites/offices offering certain holiday packages, tourist guides and books as well as questionnaires and Hotel manager interviews. For this reason questionnaires will be handed out to different people in order to gather information on what is more important for a person to choose a destination and especially from a spatial point of view, defining this way which preferences have a spatial aspect.

The most important way of gathering specific tourist preferences is to hand out a questionnaire which potential tourists have to fill in. Handing out questionnaires is the easiest, most popular and less expensive way for a researcher to gather information (Κυριαζή Ν., 2005). However, at the same time information from tourist offices, organizations and guides will be taken into consideration as well as interviews from hotel managers.

The websites, holiday packages and tourist guides will, in most of the times, provide us with activities that tourists are mostly interested in but it is going to be general information. It will not reflect a very crisp and specific group of people. On the other hand, questionnaires will be helpful to examine the preferences of a certain group of people, which is already specified above, since we are only going to hand it to people that fulfil the restrictions of this specified group.

This whole procedure mentioned above will be very helpful to gather important information which can be, later on, analysed (Zhu, X et. al, 2005). The people who filled in questionnaires could later on be the candidates to test/use the application created in order for this application to be used also in practice. By doing so we will see if they will get any useful results about their potential proposed area for summer vacation. In addition, monitoring current vacation destinations can also be helpful in narrowing tourists preferences, as already mentioned above. Another source or information can be tourist offices as well as some web-sites which offer special packages for vacation as well as tourist guides and books related to this issue. Usually the offices can provide the combinations of attractions that tourists prefer since this is actually what they do.

2.2 SPECIFYING A TARGET GROUP

At this point we have to mention that we should focus on a certain group of people in order to make our study more concrete and accurate. In general tourist behaviour and preferences vary due to a big amount of reasons such as age, family status, nationality, religion etc. For example, when visiting a Mediterranean country, tourists from European countries would rather be 'sun-sand-sea' tourists whereas people from beyond Europe would mainly focus on cultural and historical locations. (Pearce D., Grimmeau J.P., 1985) By selecting specifically the island of Paros we already narrow the tourist group to people who want to visit this island. It is also important the fact that almost every island of Greece has its own identity and in most of the cases it attracts a certain group of people with certain interests. By investigating the preferences of those who are interested in visiting this particular island we can narrow even more the group of people that we will work with. This formation of a group will be part of the research.

It is also interesting to see how tourists fit various interests they have into a certain time-frame. It would be interesting to see and suggest also what one could visit in a certain time period according to the time that they are willing to spend in total. This time needed to visit certain placed can be a certain characteristic of the places of interest.

2.3 GATHERING DATA

When all the information about which places are of the most interest for tourists is collected, the corresponding datasets would need to be obtained or created in order to be able to process them, mostly via a dynamic multi-criteria analysis to get some results. By adding the term 'dynamic' we mean that the multi-criteria analysis will not be calculated with fixed distance values but the users can fill in their own numbers according to their personal preference. The destination preferences can be defined and analysed according to certain patterns. (Gunn, 2002) At this stage we should also investigate how our preference results fit to the physical, demographical and cultural structure of Paros. We need to know whether it is realistic to propose destination on this island according to our tourists' preferences.

This of course needs further study but some potential datasets that could be useful would include:

- hotels
- beaches
- bars
- road network (with impedance preferably)
- speed (walking, driving, maybe cycling)
- restaurants
- museums

- ancient sites
- hiking paths
- ports
- airports
- slope
- aspect

2.2 CONCEPTS AND TOOLS

In this chapter we will discuss about the possibilities and ways of finding the optimal summer vacation location on the island by applying GIS methods, and especially the method of multi-criteria decision analysis. By doing so we will be able to see the use of GIS in a particular tourism problem which has a spatial nature. However, in order to understand what we are actually doing some basic concepts need to be explained at first. We need to know what spatial analysis, Geographical Information Systems and multi-criteria analysis are.

2.2.1 SPATIAL ANALYSIS

There are many different kinds of analysing information such as statistical analysis and economic analysis. Even though these ways of analysis are very useful and popular, they cannot deal with any kind of a problem. Applying different methods and techniques to features which can be defined in space does not necessarily mean that this is spatial analysis. Spatial analysis focuses on the importance of space and it is highly depended on spatial variables in order to evaluate and explain a phenomenon. On the other hand, in a non-spatial analysis, spatial factors and information is not necessarily needed. (Κουτσόπουλος, 2005). A good example can be taken from urban geography. If the population of an area expands then it is natural that the number of shops will also increase since the demand of goods will rise. Even though there is a spatial component in this example (location of the shops), the analysis is not spatial. On the other hand if we would say that the number of the shops depends on how far away they are from a city centre or from whether there are other shops in the area then such an analysis would rather be a spatial analysis.

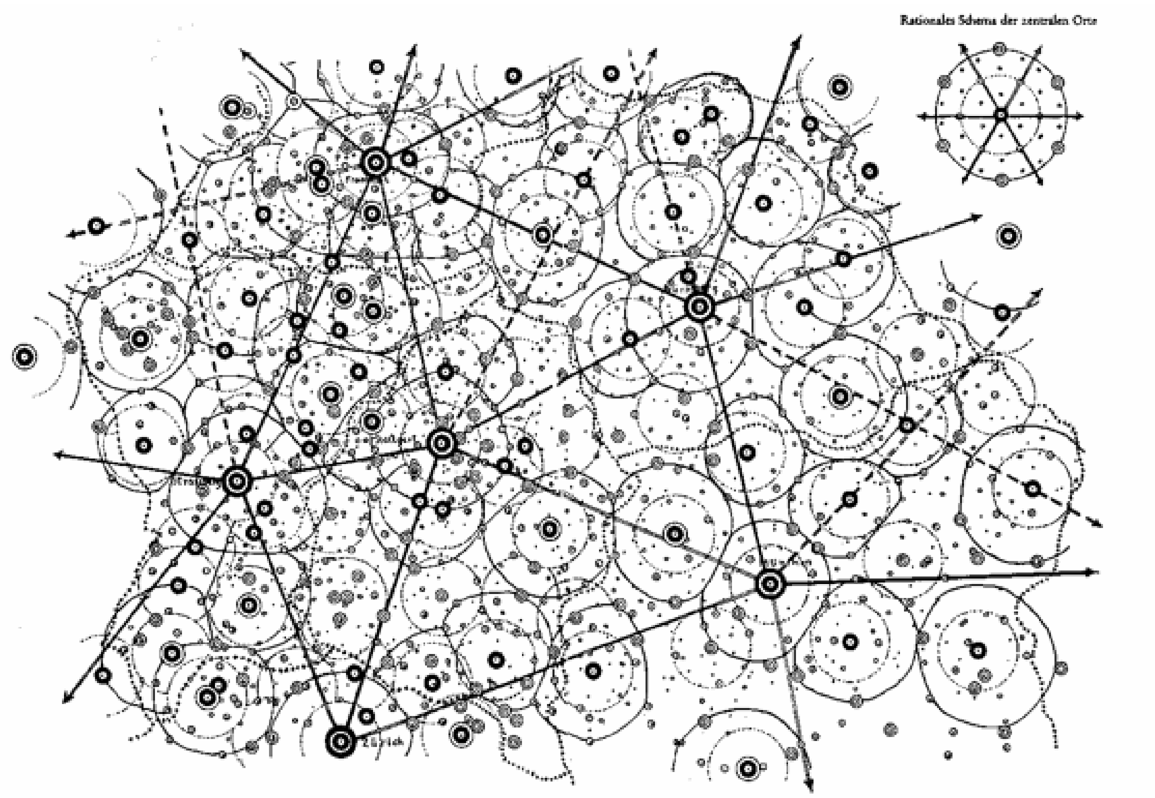

Figure 2.2: An example of spatial analysis
Source: URBAGRAM.

Figure 2.2 shows a system of the central cities in southern Germany. This figure shows the concept of the spatial distribution of the cities in the area based on the Walter Christaller's model. In this model, smaller cities are distributed in a circular way around bigger ones.

Our problem in this research also includes spatial analysis and this is going to be well seen in the next chapters where the actual analysis will take place. When calculating the optimal areas for summer vacation, this includes an analysis of spatial nature. The best example is that we have to check how close certain destinations are to each other so that one can manage to visit all of them in a certain amount of time or distance that is desired from the potential tourist. Another example of needed spatial analysis for our research would be terrain analysis. This can apply to various cases such as estimation of desired hilliness and having a nice view from a certain location. But these issues will be discussed more on a later stage.

2.2.2 VECTOR OR RASTER?

In many projects that include representation of reality in a GI System, there is always the question of whether one should use raster or vector data for their analysis. The same way we also have to choose what form is best to use in our study case. But before deciding we need to examine what the differences, advantages and disadvantages are between these two ways of representing data.

First of all we need to explain what the main differences between vector and raster data are. In vector formats every entity basically consists of nodes and arcs. These nodes and arcs in the end form the final shape of the features. For example, point features are represented with single nodes, lines with a series of nodes connected to each other with arcs and polygons are the same as lines with the difference that the lines are closed (meaning that the start of a line is the same point with the end of it). In vector data, the coordinate information is stored within the properties of the points. Figure 2.3 shows a vector representation of some simple point, line and polygon features compared to a raster representation of the same features.

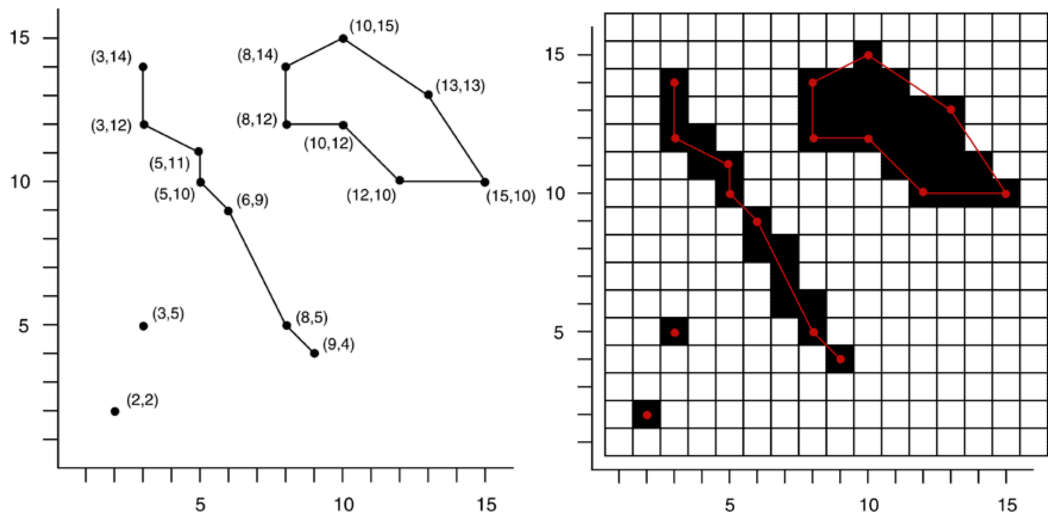

Figure 2.3: Vector (left) and Raster (right) representation.
Source: E-Education Institute.

On the other hand, raster forms represent data using tiles which are called pixels. These pixels can be of various sizes and shapes including square, hexagonal and triangular ones. However, in most of the cases one finds square shaped pixels. All three kinds of data (point, line and polygon features) are represented using pixels. For example, the simplest feature type, which is a point, is represented with a single pixel. Lines and polygons are represented by a sequence of pixels as shown in figure 2.3.

As we can see there are differences in representing data according to whether we choose vector or raster forms. The first thing that one can notice from looking at figure 2.3 is that when using the same level of resolution, the features are represented clearer and in a more crisp way when using vector forms. But this is only a first observation. Table 2.1 shows some more detailed advantages and disadvantages of vector and raster representations.

DATA FORM	**ADVANTAGES**	**DISANTVANTAGES**
Raster	Easier creation from image data	Must predefine spatial resolution
	Easy to overlay	Large amounts of storage space
	Efficient storage for dense data	Inefficient with sparse data
	Efficiently represents continuous and dense data	Deals poorly with linear data
	Easy to overlay	Hard to create networks
Vector	High resolution	Complex data structures
	Efficient storage of sparse data	Manipulations need sophisticated algorithms
	Coherent structure	Processing can take a long time
	Features can be presented without generalization	Inefficient storage of dense data
	Easier to work with networks	Not sufficient representation of continuous data

Table 2.1: Advantages of raster and vector models.
Source: Κουτσόπουλος Κ., 2005; Porter J.H.

As we can see form the above table, there are a few differences between vector and raster models. In our study we will choose Vector models but the explanation for this decision is going to be stated at the end of this chapter where all the conclusions will be listed.

2.4 OUR DISTANCE BASED APPROACH

After gathering all the information necessary for our study, we can then form the important criteria which, in combination, will provide us with the answer to the question what could make a person to decide to choose/reject an area for visiting while on vacation. This information, together with gathered literature information will form the basic criteria which will be used on a GI System to analyse and conclude which areas are most suitable. While executing this study we will also see at this stage how the initial tourist preferences can really influence the result of the suggesting areas for tourism. Based on the individual preferences, which a lot of time can also be conflicting, we can conclude into an overall evaluation of the suitable areas (Zhu, X et. al, 2005). Of course, this investigation is going to take place on a certain case study which will be the Greek island of Paros.

The method that is going to be used for calculating optimal locations is going to be based on spatial analysis of distance factors specified from the users. In more detail, there are going to be buffers around certain locations based on a preferred set distance. The combination of these buffers will provide us with a result of areas which are within the specified distances or not. The results can be used to decide upon whether a certain location would be the most preferred among others, to provide a hierarchical list of suggested options or to simply define if a certain option is considered to be acceptable or not according to some criteria that have been defined ahead of conducting the analysis. Since we are going to use the combination of distances to various locations this looks like a combination of different distance criteria and therefore it is similar to MCA problems. All MCA methods make use of some criteria and human judgement is also required (Dodgson J.S, et al., 2009). The MCA method for solving problems can also have a spatial input/ output. There are quite some spatial problems that used this method such as in the case of flood hazard zoning (Fernández D.S, Lutz M.A., 2010), landfill sites selection (Geneletti D., 2010), finding post-fire forest resilience (Arianoutsou M., 2010) and so on. In our case the MCA method will be based on spatial criteria which will be formed by the preferred maximum distances to certain places of interest within Paros island. We will build a model which will do the core calculations for estimating optimal vacation locations according to the distance criteria specified by the users.

The output of this research will give us some information on which areas of the island would be the ideal destination for a person or family. The results of such a research would look somehow like the results of Figure 2.1. Every input for the final analysis will be service areas calculated based on specified distances to different locations. The combination of these areas will give us the final result where certain areas will be coloured according to the number of locations which are within the specified distances.

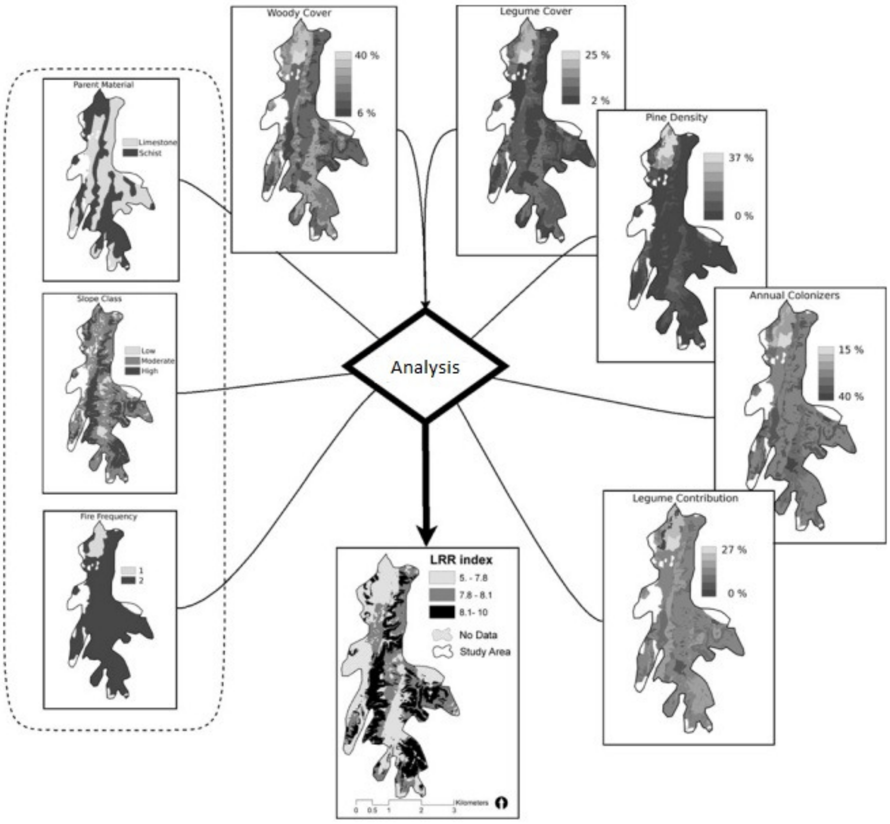

Figure 2.4: Multi-criteria analysis in finding post-fire forest resilience
Source: M.Arianoutsou, 2010

After going through the procedure of calculating the optimal places for tourism mainly using a combination of distance buffers around certain places of interest.

In general, two kinds of results should be expected out of this research according to the two main questions of this thesis. First a list of tourist preferences, which will show us what are the most important locations for them. After having specified a certain group of people that we want to examine and having checked various sources of preference listings (questionnaires, websites, tourist offices/packages) we must conclude on what is more important for a tourist when going for summer vacation.

In addition, we have to present in what degree the tourist preferences fit to our special case study area, the island of Paros in Greece. Because it is not enough to know what people want. We need to make sure that our target area has to offer what tourists want. It would make no sense if we started calculating suitable destination areas if our general area of research does not have to offer what tourists ask for.

Last but not least, there is going to be an application as an output of this study. In this application every individual would be able to set in detail the importance of their preferences, taking as a result a map which will show them which areas or Paros are suitable for them. It is going to be an application that will bring rather customised results and that is because every user would be able to set for instance the exact distance of museums from a hotel, or the exact degree of slope that they would like to have around them for whatever reason. As already explained previously, this application will strongly focus on the spatial component of tourists' preferences and more precisely on the combination of distances from certain locations. As a result the user will come up with a proposal of areas which are in a specified distance from different places of interest. The main way that this is going to be achieved is to create distance related buffers around places of interest. This application will be built mainly by using the model builder of ESRI's ArcGIS software package.

3. DISTANCE CONCEPTS

As already mentioned in paragraph 2.2, distance to certain locations is very important for tourists when they have to decide upon whether a certain place is suitable for their summer vacations. For this reason distance will play a critical role in this research. At this point we will explain the different meanings of distance according to C. Cooper (2008) that can be part of our case.

First of all there is the physical distance. This is called the Euclidean distance and it refers to how close or far an entity is from another entity in a distance unit like meters. By specifying this direct distance of a certain place from the location that one is at the moment, it is possible to estimate if that is accessible within a certain amount of time. This estimation however, also depends on other factors which also shift the meaning of distance in the end. An example is the time distance. It is calculated based on the amount of time one needs to go from one spot to another. If we assume that there are 3 points in space A, B and C with specified distance units (figure 3.1) and that we are located in point A, then we can clearly see that as an absolute distance C is much closer than B. However, if these points were cities and we knew that A is connected to B with a highway whereas the connection between A and C is only a pedestrian path, then B is suddenly closer to us from a time point of view because it would take us much less time to reach it.

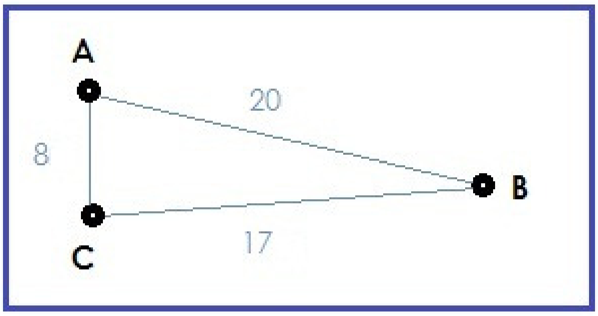

Figure 3.1: The distance between three points in space.

Another perception of distance is the so called 'Manhattan distance'. In this case, there is a certain network that connects two locations and the distance from the one location to the other can only be calculated through this network. A very good example is that a person wants to go from a certain area of a city to another by car. He would only be able to go through the road network and not straight or using any pedestrian passages.

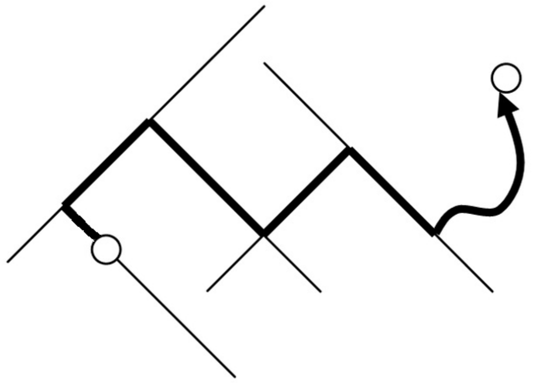

Figure 3.2: Network (Manhattan) distance.
Source: Wikibooks

As we can see in figure 3.2, in order to go from one point to another, there is no other option than following a certain predefined network.

There are many factors, natural or social, that can influence the meaning and estimation of distance between two points. For instance, there is economic distance which is calculated based on the amount of money one would have to spend to cover this distance. Social factors also influence distances because of socioeconomic differences between some locations (Cooper C, Hall C.M., 2008).

What can also influence distance measurements is slope, wind direction or temperature. It is rather obvious that if an area has an extreme relief, going uphill it will be more difficult for somebody to walk a certain distance than if he would have to walk the same distance in an area which is rather flat. In addition, one could also take into consideration the wind direction to estimate distances. If somebody would like to cycle on a road which has an orientation of northwest to southeast that is in a flat area where a lot of strong northwest winds occur, then it would take longer to go towards the northwest.

As it seems distance is a very complicated thing to calculate which can have a lot of different variables that influence the estimation. Here we only stated a few factors which are only some basic ones. In this thesis we will mainly focus on calculating physical length distances based on the Euclidean distance.

3.1 CALCULATING VECTOR DISTANCES

Since distance analysis is a main field in GIS, there are many ways to calculate distances in a GI system. In this chapter we will check these different ways so that we can decide which one is most suitable for our case. In general, the most popular methods are creating buffer zones, generating Thiessen polygons, and distance to a certain point (Κουτσόπουλος Κ., 2005). These three ways will be explained a bit more in detail in the following subchapters:

3.1.1 BUFFER ZONES

Buffer zone is a zone of a static or variable width which lays around a certain entity whether that entity is a point, a line or a polygon (Χαλκιάς Χ, 2006). The generation of these zones is a main analysis process which creates new polygons around entities that is used to specify spatial vicinity. If the entity from which distance is going to be measured is a point then a buffer zone around it could have the form of a circle or a square. The same way, a buffer zone can be created around a line or a polygon with as big width as desired (Κουτσόπουλος Κ., 2005). Figure 3.3 shows some possibilities of creating buffer zones around point, line and polygon features.

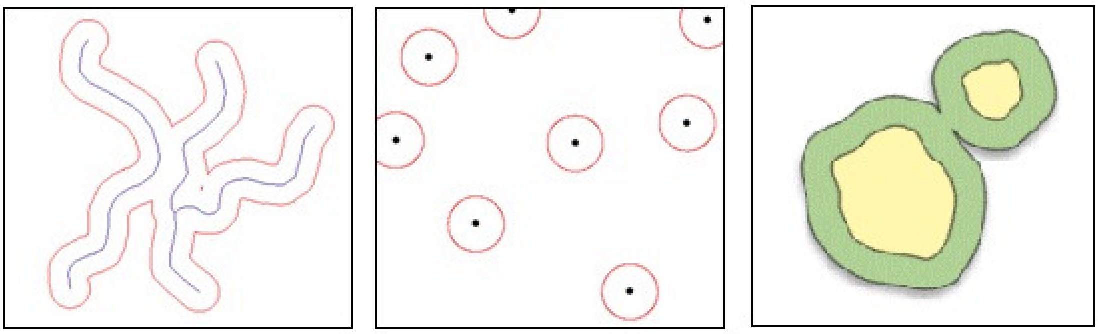

Figure 3.3: Buffer zones.
Source: ESRI webhelp.

3.1.2 THIESSEN POLYGONS

Thiessen polygons, also named Voronoi diagrams, are polygons around points which show the area of influence of each point. They are formed in such a way that the polygon which includes a point indicates the service area of the point. In other words, every part of a thiessen polygon is closest to the main point of the area than to any other point. Thiessen polygons are very useful when it is needed to divide an area to smaller sub-areas. These areas will cover the whole study area unlike buffer zones where the generated polygon is present only around the point. This of course also depends on the width of the buffer.

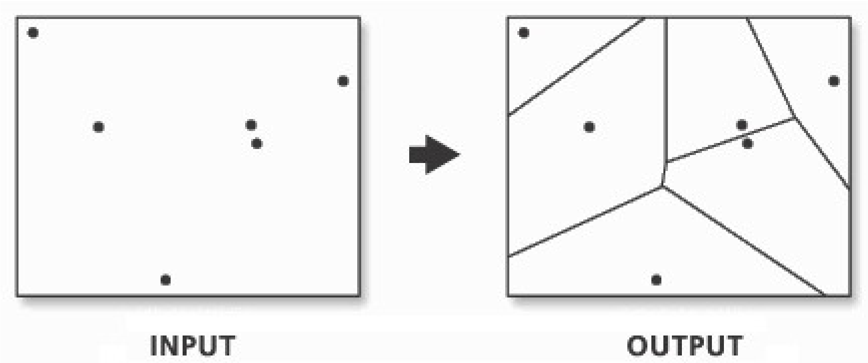

Figure 3.4: Thiessen polygons
Source: ESRI webhelp.

It is also possible to assign weights when creating thiessen polygons. There can be cases where a point is more important than another point according to its properties and of course depending on the needs of a study. In that case the shape of the thiessen polygons is adjusted according to the weights (Dong P., 2008).

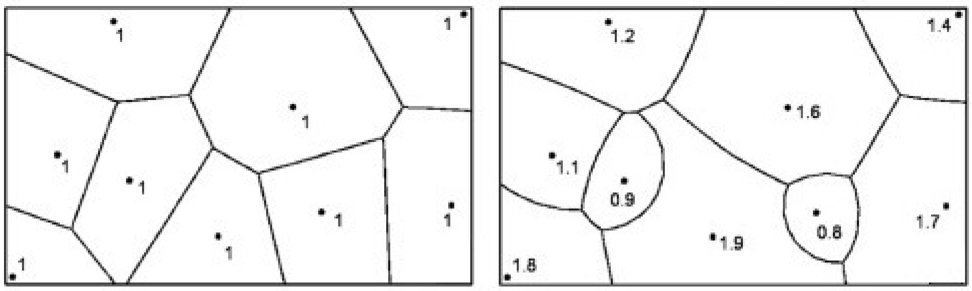

Figure 3.5: Non-weighted (left) and weighted (right) thiessen polygons
Source: Dong P., 2008.

As we can see in figure 3.5, the formation of the thiessen polygons changes a lot as we add some weight to the points that influence the generation of the polygons.

3.1.3 DISTANCE TO A POINT

In this very important method of calculating distances, the distance between a certain point and all other points within a specified area is calculated. These points could be part of the same or a different thematic layer from the one that the main point is in (Κουτσόπουλος Κ., 2005).

3.2 CONCLUSIONS

In this chapter we saw a few basic concepts of GIS as well as some tools that can be used for our case. There are two decisions that we have to take now before we proceed with building the application that has already been mentioned in the previous chapters.

First of all we conclude that for the level of this application, it is preferable to choose on using a vector model. The first, and simplest, reason for doing so is because we digitized the data we needed already, and the easiest way to do this by creating the layers needed in a vector form. Besides, the base scale of digitizing was 1:40.000 which makes it hard to transform our data to a raster format having a decent resolution level. Having in mind table 6.1 there are also a few more reasons why we choose vector format. These are because we mostly need to deal with sparse data and not with continuous. This means that we only want to have some point data and some areas around them and not a whole coverage of the area that a raster format can only provide. We basically only have point features which show the location of different places of interest and lines showing the shoreline of Paros island and the road network of the island. If we would use raster format for these layers, we would have bigger files which would not really be used in all of their extent. In addition, if we would like to take into consideration the road network when we will calculate distances (Manhattan distance calculation method) it would be rather complicated to use this method on raster-based datasets as mentioned from Porter J.H. Finally, since we will at first only calculate simple Euclidean distances which do not need any additional stored information, raster would just be a lot of unnecessary data stored for no reason.

Furthermore, since we only calculate distances and not any other kind of an impedance, raster datasets are not so useful. It is true that raster datasets are very useful when one wants to store information about temporal or financial impedance because these values do not measure something physical. In that case the impedance values would only estimate the cost of crossing a certain cell. Therefore, since our study only focuses on distances and especially those along a specified network, raster is not the best way to sufficiently represent it.

The second conclusion we have to make is the method we are going to use for calculating distances. Since this is going to be a rather basic application, we will mostly calculate Thiessen polygons and Euclidean distances (Buffer zones). The buffer zones will only be calculated based on absolute distances.

4. REFERENCE GROUP

In order to be able to start with our research we need to first make sure that we will collect the basic data that we need. In our case, that would be the preferences of the tourists who are planning on visiting Paros island for summer vacation[1]. These preferences will be the core information of our further study as they are going to be the base of forming the criteria of our analysis. For this reason we need to pay a lot of attention on what information we have to gather exactly and from where. Since we are going to make some suggestions on optimal summer vacation locations, the tourists' preferences are going to be the base of this calculation.

Figure 4.1: Going on vacation
Source: http://collegelifestyles.org

Gathering information about tourist preferences may sound and easy task to do but in reality it is a rather complicated issue. First of all, we have to specify a crisp group of people that we are going to gather information from. It is not possible to take into consideration the preferences of every single kind of tourist, especially within a Master thesis, because then our target group would be enormous. There are many kinds of tourists who travel for different reasons (Wickens, 2002). That could include cultural tourists, leisure tourists, ecology tourists, and so on. In addition, there could be different groups from a social status point of view like families, retired people, couples etc. It is obvious that all these different groups of people have different priorities and therefore different interests while on vacation. Since not all groups fit to all

[1] By summer vacation we mean the middle and high tourist season, according to hotels information. This period is on average the period between the 1st of June until the 31st of September.

destinations, in this work we are going to narrow down to a specific group of tourists that is going to be dealt with.

4.1 TOURIST TYPES

In general, one can divide tourists, who come to Greece for summer vacation, into groups according to their preferences. From a study made in Chalkidiki in Greece (Wickens E., 2002), there were five tourist types based on tourists' holiday preferences. However, in this study we will mostly focus on three of those types which most fit to our tourist group specified at the previous chapter; the 'Cultural Heritage Type', the 'Raver Type' and the 'Heliolatrous Type'.

In the first group the tourists focus (as revealed from the type's name) rather on cultural interests such as museums, monasteries and so on. In addition, according to a participant in Wicken's study, food, animal/ plant life, Greek villages, and water sports were also important.

The Raver type of a tourist is characterised by their interest in sensual and hedonistic pleasures. These tourists are only searching for ways to just have fun and be entertained during their stay. 'Ravers are on a quest for thrills rather than for authenticity or Greekness' states Wickens. Therefore those people are mostly interested in chilling on beaches, going to bars and night clubs.

The third category, the Heliolatrous type includes the tourists who have as a priority to have sun during their vacation regardless where they actually go. (Wickens E., 2002; Gibson H., Yannakis A., 2002) The word 'heliolatrous' comes from a combination of two Greek words which mean 'to love the sun'. This kind of tourists would rather go to places where they can get as much sun as possible, especially those who come from countries that are cloudy most of the time like England and Holland. The most popular place for a tourist to go and 'gather' some sun is the beaches. For this type of tourists the most important factors of a summer holiday is going to beaches and restaurants (Wickens E., 2002)

4.2 FORMING A TEST GROUP

As explained above, we need to specify a certain kind of a tourist group from which we will extract the tourist preferences. Only by concluding on a certain group of people we can continue with our study. We have to mention here that there are a lot of attributes on which a group can be specified (Decrop A., Snelders D., 2005) and therefore, we can practically form a very big amount of different tourist groups with diverse preferences. However, it is not possible to take into account so many tourist groups as this could even be an independent topic that one can work on. In this study we will take into account a few important groups of tourists and then we will extract their preferences.

4.2.1 ORIGIN

In many cases tourists differ according to where they come from. It is important to be able to realise the origin of the tourists because this also influences their behaviour and their preferences.

In general, there are three main categories of tourists: those who come from the same country, in this case Greece, those from Europe and those from the rest of the world. (Pearce D. G, Grimmeau J.P., 1985). This grouping was for a study of Spanish tourism but being also a Mediterranean country it can be compared to Greece. In the following figure we can see how European tourists and tourists from the same country overrule the tourists from the rest of the world.

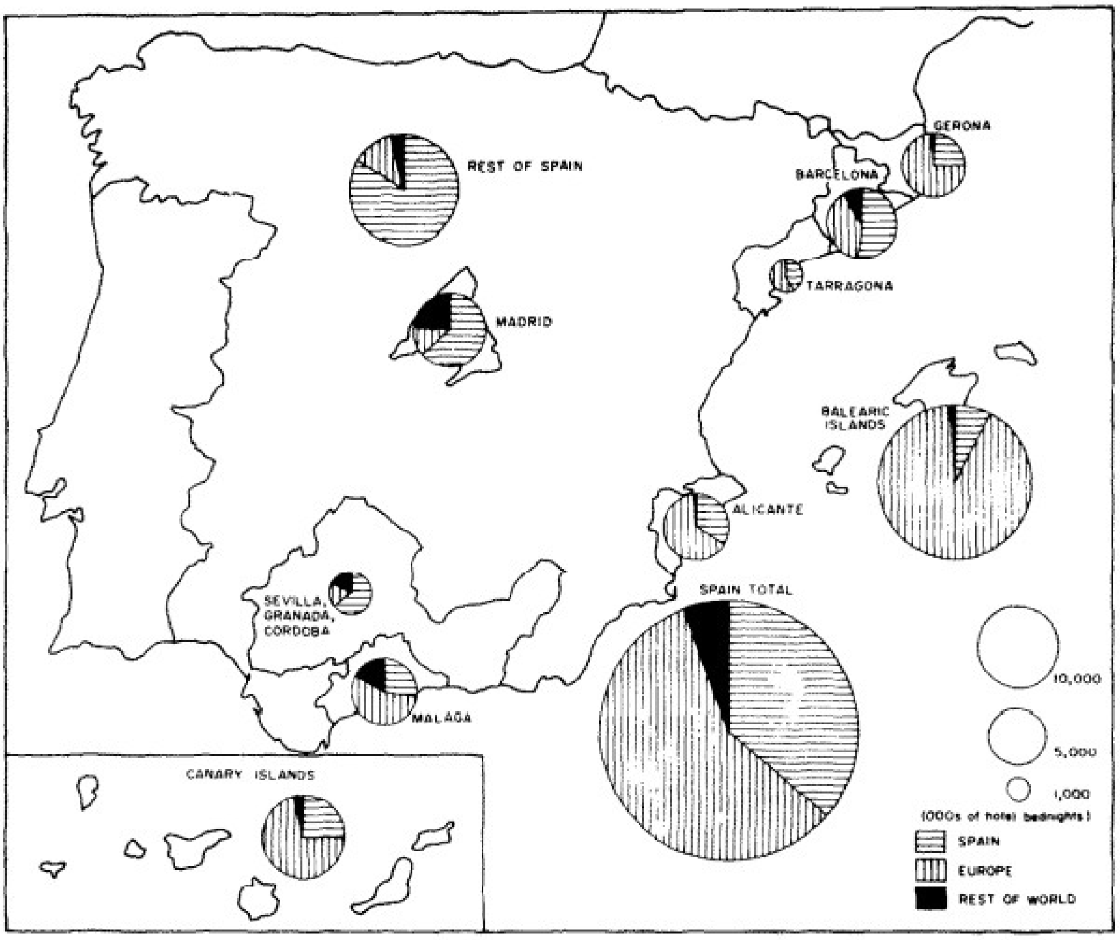

Figure 4.2: Tourist distribution in Spain.
Source: Pearce D. G, Grimmeau J.P.

In figure 4.2 we can see that in Spain, most of the tourists come from the country itself and from European countries. It is obvious that the rest of the world only has a small share, with the exception of Madrid. Now if we have a closer look at the Mediterranean side of the country, which is more comparable to Greece, most of the tourists are from Europe.

According to Wittand and Martin, the most important tourism generating countries for Greece are mostly European countries and the USA. (Witt S.F. and Martin C.A., 1987) The following table is also going to prove that, between the forty first countries of origin of the Greek tourists, there is an overwhelming majority of European tourists.

in thousands	R00	R96	96	97	98	99	00	in thousands	R00	R96	96	97	98	99	00
G.Britain	1	2	1.688	1.712	2.044	2.433	2.772	Cyprus	21	22	86	131	127	139	135
Germany	2	1	1.908	1.995	2.136	2.450	2.395	Turkey	22	28	47	45	70	80	134
Italy	3	3	491	533	660	746	823	Israel	23	24	74	82	89	155	116
Albania	4	13	191	299	586	673	717	Spain	24	23	80	71	97	99	115
Netherlands	5	5	452	464	548	617	655	Romenia	25	27	51	52	63	73	99
France	6	4	463	427	486	546	602	Japan	26	21	87	85	87	84	78
Sweden	7	6	448	472	468	469	486	Russian Fed.	27	14	180	201	127	105	102
Austria	8	7	360	388	450	502	475	Slovakia	28	30	28	30	55	46	60
Denmark	9	8	299	344	292	336	339	Ireland	29	26	54	45	44	49	59
Belg. - Lux.	10	12	213	229	274	333	332	Australia	30	25	54	47	43	50	58
Switzerland	11	9	256	296	289	308	322	Canada	31	29	44	48	50	52	56
Norway	12	16	158	160	226	269	314	Ukraine	32	34	13	27	28	33	nd
Bulgaria	13	17	155	182	197	203	240	Egy.-Sud.	33	31	20	19	18	24	37
R.Macedonia	14	---	---	378	87	128	234	Leb.-Syr.	34	32	18	16	18	18	23
U.S.A.	15	11	222	240	219	229	219	R.S.A.	35	36	9	8	7	10	12
Czechs R.	16	15	177	174	178	174	197	Portugal	36	33	16	13	14	21	9
Jugoslavia	17	10	231	199	202	103	159	Argentina	37	37	8	4	4	5	5
Finland	18	18	121	148	149	189	156	Brazil	38	35	10	8	6	4	4
Poland	19	20	89	102	123	115	154	Mexico	39	39	3	2	2	4	4
Hungary	20	19	114	145	186	123	148	Iran	40	38	5	4	4	4	3
E.U.	-	-	6.593	6.843	7.663	8.789	9.219	Total	-	-	9.233	10.070	10.916	12.164	13.095

Table 4.1: The 40 first original countries of the tourists in Greece in 2000 in thousands.
Source: www.eurogreece.net

All the research results above show that Europeans are the majority of the tourists in Greece. However this is without taking into account the Greek tourists themselves travelling within their own country for vacation. The National Statistics Organization of Greece (www.statistics.gr) as well as Eurostat (http://epp.eurostat.ec.europa.eu) state that most of the tourists of the country actually come from the same country. This can be seen in table 4.2 below.

Country of origin	Amount of tourists
Greece	6.000.000
E.U.	9.000.000
Rest of the world	4.000.000

Table 4.2: Numbers of tourism in Greece in 2000
Source: National Statistics Agency

As we can see in table 4.2, the amount of European tourists in the year 2000 was the biggest of all which means that most tourists that visit Greece are from European countries. The next biggest amount is inland tourists, meaning tourists from within Greece. Since there was no exact information about these numbers in Paros specifically, we have to assume now that the percentages are the same also in Paros.

From the above information it seems that, especially for Greece, most of the tourists are European people and theoretically we could only take into consideration only these people to our study. But since tourism is a very important economic factor for the islands of the Aegean Sea (Papanastasiou A. et. al, 2006), and since Greek tourists are also an important tourist group, we will include both Greek and European tourist groups into our research.

4.2.2 AGE DIFFERENCES

Age is also a very important indicator to categorize tourists in groups. Every person has different interests and therefore different preferences while on vacation according to their age. It is not possible that a person aged 18 has the same enthusiasm for an activity as someone at the age of 40 or 65. Since we want to narrow our tourist group down to a more manageable group we have to decide which age group we are going to work with.

According to Urhausen, there are 4 main age groups of different tourist behaviour and these are people aged 15-24, 25-44, 45-54 and 65+. Most of the tourists come from the age group of 25-44 and 45-54 in all European countries and therefore also in Greece. (Urhausen J., 2008)

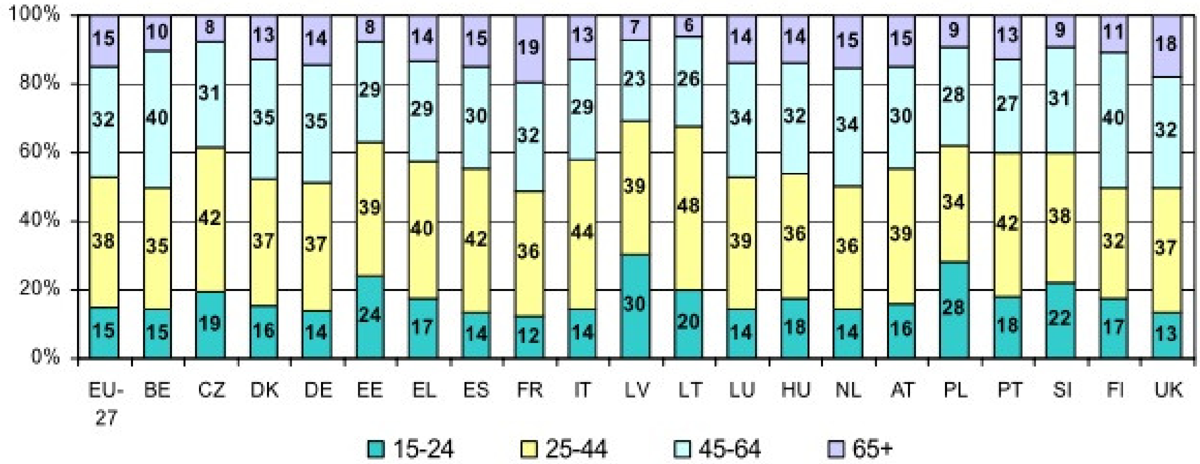

Figure 4.3: Share of each age group as a percentage of the total number of tourists in 2006
Source: Eurostat

Figure 4.3 proves that most of the tourists are aged between 25 and 64 years old. That makes sense due to the fact that this age group contains the people who are economically active and are still at an age that leaving from home to travel will not cost them as much energy as it would for a 70 or 80 year old person. Tourists aged younger than 25 years old are also not so many since, in most of the cases, they do not have their own savings to finance a trip or they are too young to travel alone.

Concluding from the above mentioned information, there are two age groups which form the biggest tourist amount. These are people aged between 25 and 44 years old and between 45 and 64 years old. So in general, the age group of people that is going to be examined will be between 25 and 65 years old. Since these people are those who form the larger groups of tourists, we should mostly focus on their opinions.

4.2.3 THE IMPACT OF GENDER

Another factor to distinguish tourists is their gender. It happens a lot of times that men have different preferences than women. Men are more interested in spending time in sport related activities whereas women prefer dancing, shopping and visiting family members. However, when it comes to vacation time, these differences become smaller and smaller and one cannot compare them to the differences when at home. (Carr. N, 1999) According to Swain, when people are on vacation, some habits that are considered to be mostly male habits also become female habits. An example is that drinking is something that mostly men would be interested in, however, when on vacation women seem to be also drinking. (Swain M.B., 1995) In addition, Carr also claims that 'the behaviour of young men and women (18-34 years old) is now becoming increasingly similar, blurring the boundaries between the genders'. (Carr. N, 1999) Furthermore, the fact that age has a significant influence to the tourism destination is only partially confirmed. (Beerli A., Martín J.D., 2004) The following table also indicates how, in most of vacation activities, men do not differentiate so much from women.

Differences between the activities engaged in by young men and women tourists (% of tourists)

Frequency engaged in	Shopping for souvenirs		Taken photos		Engaged in sport/ physical activity		Sunbathing/ relaxing		Visited places of interest		Walked around resort		Visited theatre/ cinema		Eaten local food		Walked around area surrounding resort	
	Women (W)	Men (M)	W	M	W	M	W	M	W	M	W	M	W	M	W	M	W	M
Very Frequently	26	12	18	12	12	12	28	30	17	22	56	61	6	2	11	27	24	20
Frequently	20	6	20	11	4	15	18	21	21	18	23	24	5	5	17	15	32	21
Rarely	16	22	17	19	19	23	12	18	19	21	7	6	14	19	19	18	19	28
Very rarely	38	60	45	58	65	50	42	31	43	39	14	9	75	74	53	40	25	31

Table 4.3: Gender differences in tourism
Source: Carr. N., 1999

Table 4.3 shows that both men and women have more or less the same interests with only a few exceptions.

From all the above statements we would like to conclude that it doesn't really matter if we would include both genders or just one, since the difference is not so big. For this reason we are going to include and combine in our study the opinions of both male and female tourists.

4.2.4 MARITAL STATUS

The tourist characteristics can also depend on the marital status of them. According to Mahasuweerachai, 'marital status is significant in tourists' characteristics'. There are differences between people who are married, or in couples, and those who are single. The question is which of these tourist groups are more likely to travel.

In Meng and Uysal's study about natural-based resort destinations, it seems that married people are more likely to travel than single people, with a rather big difference. (Meng F., Uysal M., 2008) Having a look at the table taken from the same study, we see that the potential visitors are mainly married with a percentage of 63.7% which is much bigger than the 19.7% of the single people.

Marital Status	Percentage of potential visitors
Single	19.7
Married	63.7
Divorced/ Widowed/ Separated	13.3
Other	3.2

Table 4.4: Potential visitors
Source: Meng F., Uysal M.

As a conclusion we can state here that based on the marital status factor we will only take into consideration the tourists who are married or in couples.

4.2.5 THE INFLUENCE OF CHILDREN

Families with children form undoubtedly a special group of tourists since their preferences influence what the family is going to do during their vacation time. Children's wishes most of the times conflicts with those of the parents/ adults. According to a survey of Thornton et. al, there are activities that are more interesting for children and others that are more interesting for adults, and that is shown from the result of the time budget diagram (figure 4.4). (Thornton et. al, 1997)

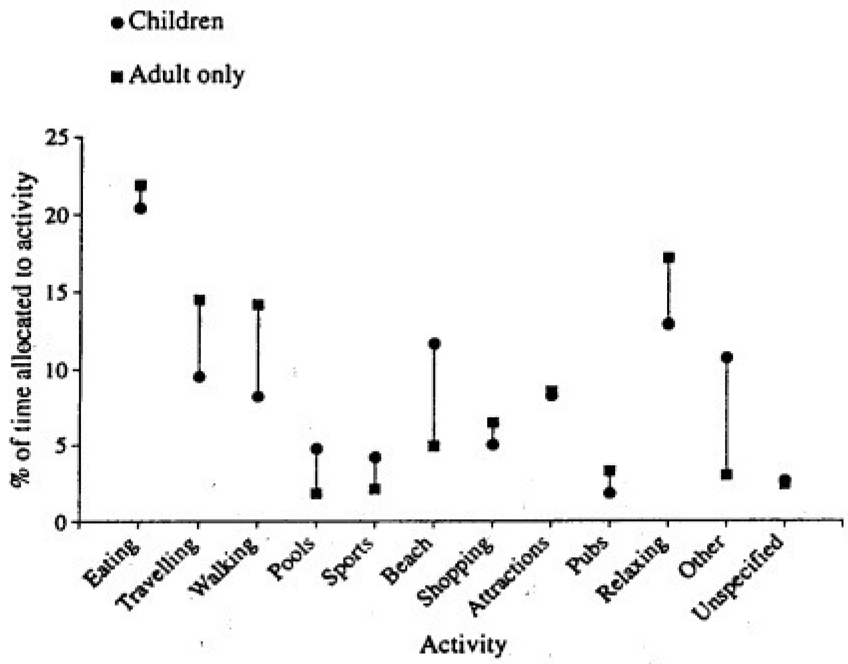

Figure 4.4: Time-budget diagram
Source: Thornton et. al

As one can see from the above figure, in many cases, children usually spend different amounts of time on several activities, and that is because they have different interests. In some of these activities, the difference of interest is rather big. For example, Travelling, walking and relaxing is something that adults mostly do whereas children prefer spending a lot of time at the beach, in pools and sports. Of course, the older the children get, the smaller the difference with adults but still their interests do not match. (Thornton et. al, 1997) Even though all the activities have a clear difference between adults and children, there is the 'other' activity which is unclear in this diagram. Unfortunately, in Thornton's work it is not mentioned which activities are included, not even in general. On top of that it seems that in these activities the difference between adults and children is noticeable. We can assume here that it might be activities which almost only children are interested in, however we cannot go any further than that.

At this point we can say that if children are included in a tourist group then they are most likely going to influence peoples' decisions and the adults preferences will be altered to fit their children's needs. That would result in a deviation of the original choices of the adults. As a result, the adults' preferences will not be real, but adjusted to those of their children. Since in this research we want to have, as much as possible, the real desires of tourists we will choose to only deal with adults and not with families who also have young children.

4.3 SELECTED TOURIST REFERENCE GROUP

After having researched which groups of people are more likely to travel to Greece, according to specific categories of grouping, we can define the focus group of this research in order to extract the most important preferences of these tourists.

Our tourist group will include people who do not travel with children, who are married or in couples (not single), are between the age of 25 and 64, and those who come from Greece or another European country. Now having this group of people as our research group we will extract the preferences of these people, which preferences are going to indicate us which are going to be our criteria four our distance-based solution of the problem.

5. EXTRACTING PREFERENCES

After having specified the group of tourists that will be taken into consideration, we can now proceed in getting to know what are in general the preferences of those tourists when they are planning their vacation. These preferences will be used to form the criteria for our distance based application. Since this thesis has a geographic component we will focus on tourist preferences which will form a spatial criterion, in particular, such as distances/proximity to certain locations, slope percentages and so on. One of the research questions is what is the most suitable place for a tourist to go on vacation. In order to do this we would generally have to go through two steps. First to find out which places (facilities, amenities) would be of interest to most of the tourists and second how important each facility is for them (Zhu, X et. al, 2005). The destination preferences can generally be defined and analysed according to certain patterns. (Gunn, 2002). In whichever way, after gathering information about tourist preferences we will be able to also decide later on which data is going to be of great use for calculating the optimal areas for tourism in Paros island.

In the following chapters we will investigate what information can be collected from four different sources and in the end we will be able to conclude on what is most important for the tourist visiting Paros.

5.1 PREFERENCE SELECTION

In order to be able to proceed to any kind of analysis, we have to narrow down and be more precise about what is most important for tourists while on summer vacation also from a spatial point of view. For this reason we collected information from different sources which are the most popular for this kind of information. We have to mention here that the questions addressed to the hotel manager, the tourist office staff as well as in the questionnaire are more or less the same so that we have homogeneous answers that can be later on compared.

5.1.1 TOURIST GUIDES AND WEBSITES

In order to collect some information about tourist preferences through the web, four different websites were visited. This selection of sites was based on finding the most popular ones. After going through a number of holiday booking websites and guides we found out that there are a few activities which are mostly present in all of these sources and which form the most important preferences of tourists. In particular, these interests are: going to the beach, visiting monuments as well as archaeological sites. Of course it is important to keep in mind that except for the activities, accommodation is equally important for a tourist since without a place to stay a tourist can basically go nowhere and therefore it is the first thing that a tourist will start looking for.

5.1.2 HOTEL MANAGER INTERVIEW

Another very important source of information is to also ask a hotel manager about their customers' preferences and what they ask for when they book a room, once they arrive at the hotel and during their stay. For this reason we contacted Ms. Kalamboka, the owner and manager of Sea View Hotel in Dryos town of Paros, in order to extract useful information. According to Ms. Kalamboka (see Appemdix 3) what tourists are mainly interested for is beaches, restaurants, super markets and bars. In addition, those who like sports they also ask for places to go for wind surfing as well as horse back-riding. Furthermore, it is important for them to have accessibility to the island travelling by airplane and therefore for them an airport needs to be preferably present.

5.1.3 TRAVEL AGENCY INFORMATION

In order to obtain additional information about tourist preferences, there was contact with two travel agencies in Thessaloniki, Polizas Travel Services and Kronos Holidays (see Appendix 4). According to their summer packages for Greek summer vacation destinations it seems that they mainly focus on visiting archaeological sites and monuments. However, Polizas Travel Services said that most summer packages have a rather free programme where the office organizes the transportation and accommodation issues and the rest of the activities are upon the tourists' decision. Both travel agencies claim that the most important activities are visiting archaeological sites, monuments and thematic parks.

5.1.4 QUESTIONNAIRE

After having gone through various sources for tourist preference information, as seen above, we check peoples' actual summer vacation preferences. For this reason we set up a simple questionnaire (see Appendix 2) asking 38 individuals to put in order the most important activities for them as well as to add activities that they personally think that are important which were not included in the list. The people to which this questionnaire was handed to, were strictly those who meet the criteria of the tourist group which was formed in the previous chapter and it was given to them randomly by sending emails or by interviewing them in person. This means people who do not travel with children, who are from the European Union and are aged between 25 and 64 years. The fixed list of the questionnaire included the options of: going to the beach, going to restaurants or taverns, visiting monuments or monasteries, doing sports (including water sports), visiting archaeological sites and going to bars or cafeterias. According to the results of this questionnaire, the importance of the suggested activities is as shown below:

1. Going to the beach
2. Going to a restaurant/ tavern
3. Visiting archaeological sites
4. Visiting monuments/ monasteries
5. Going to bars/ cafeterias
6. Doing sports

In addition to this order of important activities, 11 out of the 38 candidates also filled in some additional activities that are of importance to them when on summer vacation. It is expected that for every person the preferences can be very different but we can show here which were the additional activities that most of the candidates filled in. These are mainly visiting natural parks, going for shopping and also visiting museums.

5.2 THE FACTOR OF DISTANCE

As already mentioned previously, the final purpose of this thesis is to provide tourists with customized maps which will show them where they should go for summer vacation according to personal preferences. The innovative part here is that the results will not be location driven but distance from certain places is going to be a major criterion which will contribute in estimating the right place to go to.

According to Tobler, 'The first law of Geography is that everything is related to everything else, but near things are more related than distant things' (Tobler W.R., 2004). Distance is a very important factor also in terms of tourism since it is a very explanatory variable for tourism production. Travel behaviour can be strongly related to many forms of distance, whether it is spatial, economic, network, cultural or social distance (Hall C.M., 2005). But usually most of the tourist decisions are taken in order to minimize, as much a possible, the frictional effects of distance (Cooper C, Hall C.M., 2008).

Distance plays a very important role in tourism and for this reason we are going to include it as a major factor which influences the decision of tourists on the most suitable holiday destination for them.

5.3 FINAL CRITERIA - RESULTS

Until this point we have managed to extract some preferences of tourists through a number of sources. As explained at the beginning this whole procedure was to be able to form some criteria on which tourists decide upon their summer holidays. The collection of this information will eventually help us decide upon which datasets we will also have to obtain or create in order to use them in our model later on.

After our research about tourist preferences we have to conclude to a number of those which appear most so that we can decide which are going to be part of our model and which not. Taking into consideration the suggestions from all the above ways of collecting information we will conclude that the activities that will be part of this research are:

- Beaches
- Restaurants
- Archaeological sites
- Natural parks

These are the four activities that we will take into consideration in addition to some basic features such as hotels and transportation locations. In addition, the very important criterion of distance to the locations which include these activities will also be taken into consideration. This will give the innovative spatial component to this study and will make it different than the already existing ones. All the data needed, will be collected from various sources and are going to be the main criteria input in the application but this will be better shown in the next chapter. The step that is going to follow is to check if Paros island can offer what the tourists like for summer holidays but this will be analysed in the following chapter.

6. PAROS ISLAND

It is very good to know what tourists want in general, but what if our selected case study of Paros is not a suitable example? It would be a waste of time to work on a place which cannot satisfy potential tourists. In order to recommend suitable locations for vacation, the activities that the area offers have to match the traveller's preferences (Huang Y., Bian L.; 2009) . For this reason we have to make sure that Paros island can fulfil the needs and expectations of tourist before going more in depth in this thesis.

In this chapter we will say a few things about Paros island and examine what the island has to offer to potential tourists. This information will help us conclude on whether this island is suitable for summer vacation, having in mind what our formed group of tourists are most interested in. It is very important to check not only what a tourist can view and visit, but also where the interesting, for the tourists, are located (García-Grespo A. et al.; 2009). According to a few organisations, we found out about the existence of some utilities in Paros. The main source of our information is the Municipality of Paros (www.paros.gr), the Greek Yellow Pages (www.xo.gr), and also one of the biggest booking online systems of Greece (www.pamediakopes.gr).

6.1 GENERAL CHARACTERISTICS OF PAROS

Paros is an island located at the Aegean Sea in Greece and it is part of the Cyclades island constellation. It is situated more or less in the middle of the constellation and is the 20th biggest island of the Cyclades, with an area of 197,2 m^2 and a coastline of 118 km. Its distance from the mainland of the country is about 114 km. As for the relief of the island we have to mention that it is in general a rather hilly island with the highest point of it having an altitude of 771 m (Χαλκιάς Χ., 2002; Μπομπότη, 2008).

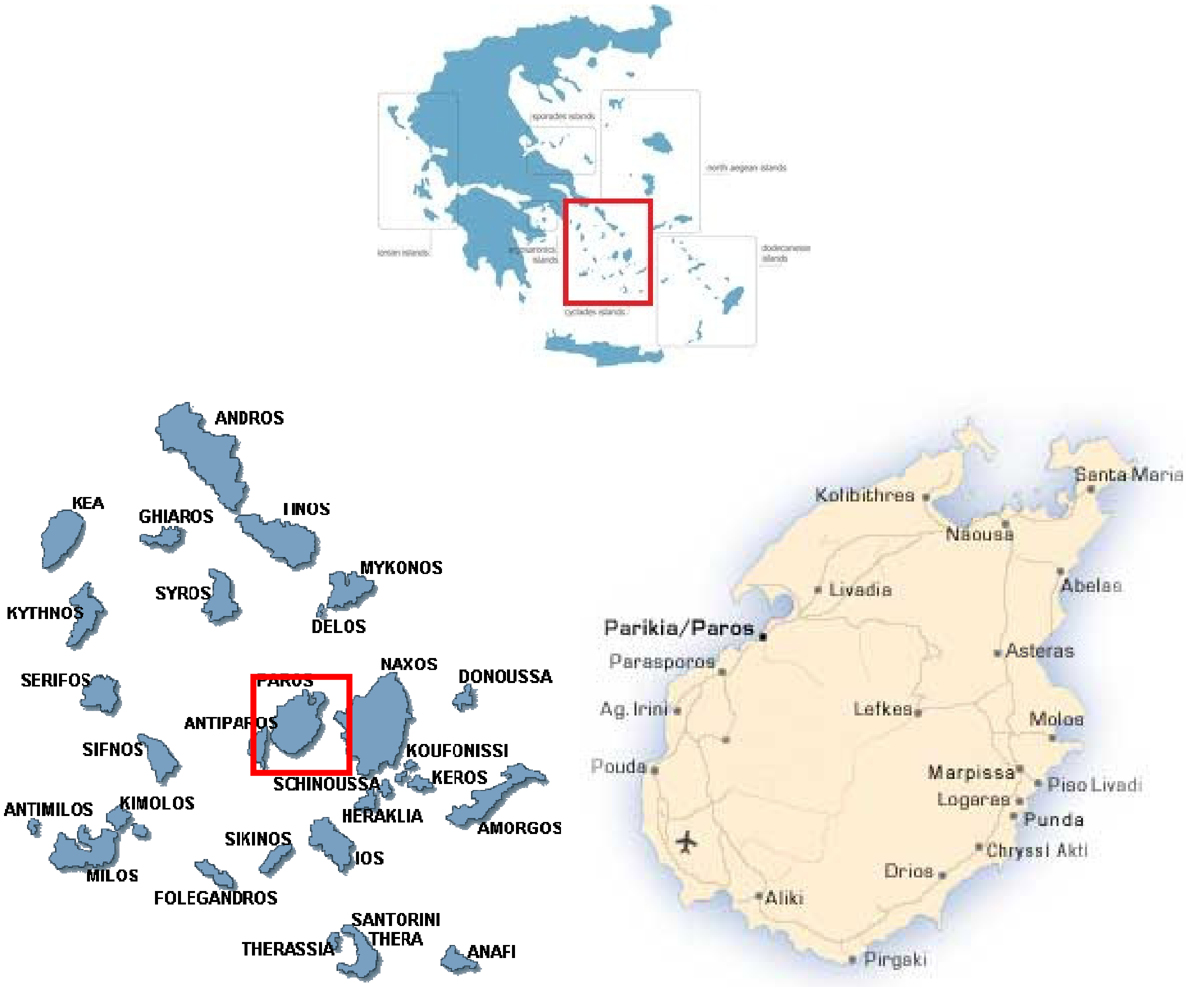

Figure 6.1: Greece, the Cyclades constellation and Paros island.
Source: www.yacht-charters.gr, www.cyclades.info, http://www.parosisland.com/

6.2 WHAT PAROS HAS TO OFFER

In order to really check whether Paros is a suitable island for our tourist group we cave to compare the tourists' general summer vacation preferences to what Paros can offer. It would make no sense to create an application for an area that does not fulfill the main needs of the tourists. We have to make sure that our case study selection is suitable for this research. For this reason we are now going to examine if one can find and do the things suggested above while visiting Paros. All the maps shown in the next pages are in a scale of approximately 1:90.000.

6.2.1 BEACHES

According to the results of our previous preference research, the most important thing for a tourist for summer vacation is to have beaches. Fortunately, Paros is an island which has many beaches which are also considered to be very good and clean. Four of the island's beaches were even awarded with the 'Blue Flag'[2] in 2009 which is an eco-label award from the independent non-profit organisation Foundation for Environmental Education (www.blueflag.org, www.eepf.gr, 8.4.2011).

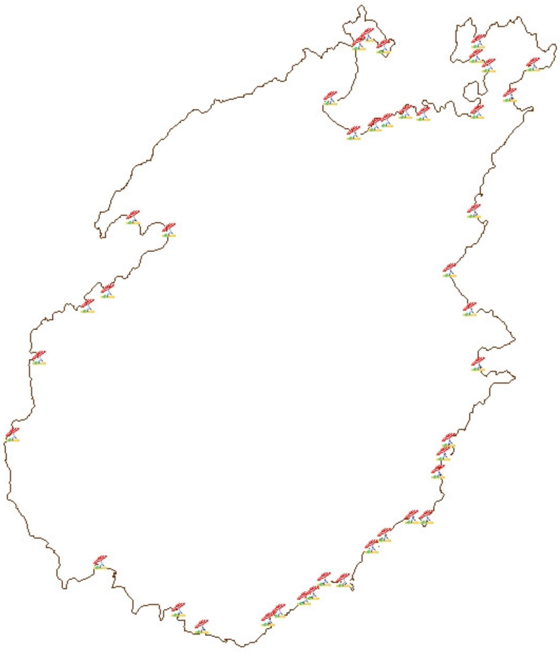

Figure 6.2: The beaches of Paros
Source: Anavasi

As one can see in figure 6.2, Paros has many beaches but most of them are gathered to the north of the island as well as in the southwest coast of the island. Therefore, we now make sure that beaches, which are the most important reason to go for summer vacation, are definitely present in Paros.

6.2.2 RESTAURANTS

The second most important interest of summer tourists is to have a good restaurant or a tavern at the area of their destination. Apparently it is important for people to have

[2] The Blue Flag works towards sustainable development at beaches/marinas through strict criteria dealing with water quality, environmental education and information, environmental management, and safety and other services

good food somewhere close while they are on vacation. This makes sense since eating is a major need of a person throughout the day. Paros, like every other island, also has a few Greek restaurants/ taverns scattered throughout the island. The following map shows the presence of restaurants in different places of Paros.

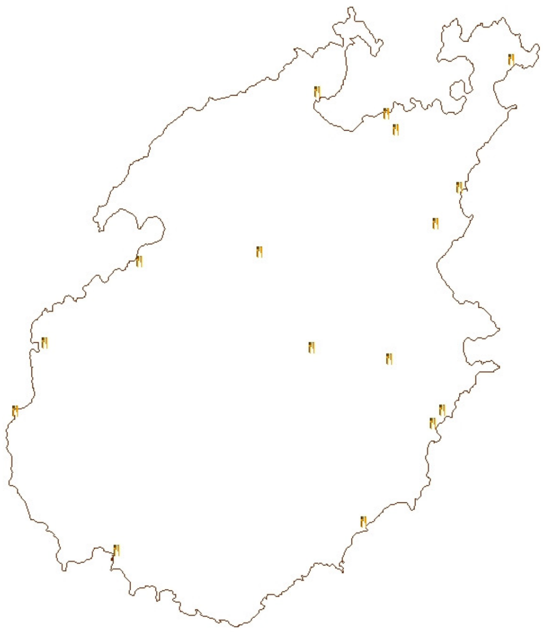

Figure 6.3: Restaurant presence in Paros.
Source: Χρυσός Οδηγός

In figure 6.3 there is an overview of the places where one can find a restaurant. Not every single restaurant is shown but whether restaurants exist in a certain town or not.

6.2.3 ARCHAEOLOGICAL SITES

Greece is known for its archaeological history and therefore for the numerous sites spread around the country. For this reason, many tourists, both Greek and international, are interested in visiting these sites during their holidays. Like almost every place in Greece, Paros also has to offer a couple of interesting archaeological sites including the ancient mines in Marathi. The following map shows the archaeological sites that are in Paros island.

Figure 6.4: Archaeological sites of Paros
Source: Municipality of Paros

Being only a rather small island, Paros does not have so many archaeological sites. However, for those who like visiting such sites there is the possibility to do so. Since the results of our questionnaires show that most of the tourists are interested in archaeological sites then it was necessary to include them to our research.

6.2.4 NATURAL PARKS

Paros island also has to offer a couple of natural parks for those who are fans of nature and have as a high preference to be able to visit an area of natural beauty. Even though Cyclades is an island constellation which is considered to have one the driest climates of Greece, Paros has to offer two parks of interest. Firstly the Butterfly park, at the southeast part of the island, and second the Environmental and Cultural Park, which is located in the north of Paros. Figure 6.5 shows the exact location of these parks.

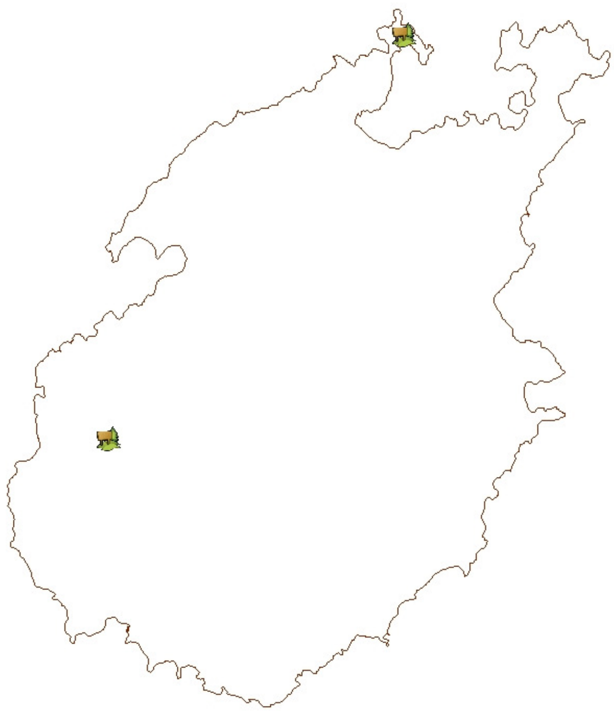

Figure 6.5: Natural Parks in Paros

6.3 ADDITIONAL IMPORTANT LOCATIONS

Even though we have already examined whether Paros island fulfils the expectations of most of the tourists for summer vacation, we need to also check if some other important places are also present. For example, where one can find a hotel and how and where a tourist can arrive to Paros.

6.3.1 HOTEL PRESENCE

It is very important to know where there are hotels on an island. This is because there is almost no tourist that will not want to have a place to spend the nights at during their holidays and it can be that they even plan their whole holidays based on where they will stay. As it seems in figure 6.6, there are some hotels around the island. We have to mention here that on this map not every single hotel is shown. It is rather an indication of whether a hotel is present at a certain town. For instance in a big town it is obvious that more that one hotel is present. For reasons of convenience, that we will see later on the study, we decided to work this way and not taking into consideration every single hotel on the island.

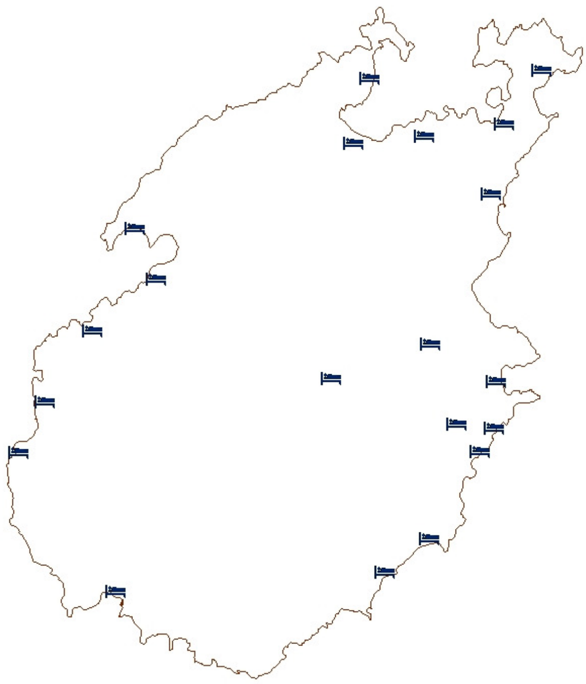

Figure 6.6: Hotels in Paros
Source: Anavasi

6.3.2 ROAD NETWORK

What is also additionally important is the general accessibility of Paros. We have to check how accessible the island is by looking on the ways and times one needs to reach/abandon Paros. According to Rigas, there are advantages and disadvantages to both modes of transportation (air and sea) and therefore there are not big differences in choosing between them (Rigas K., 2009).

In our case, Paros is an island that is considered to be a big island in the Cyclades and therefore it hosts both a port and an airport. Taking the ferry from Pireaus (port of Athens) to Paros can take from 2.5 to 5 hours (Blue Star Ferries, Hellenic Seaways), depending on the type of vessel, while flying there takes 45 minutes (Olympic air). Figure 6.7 shows the locations of the port and the airport of Paros.

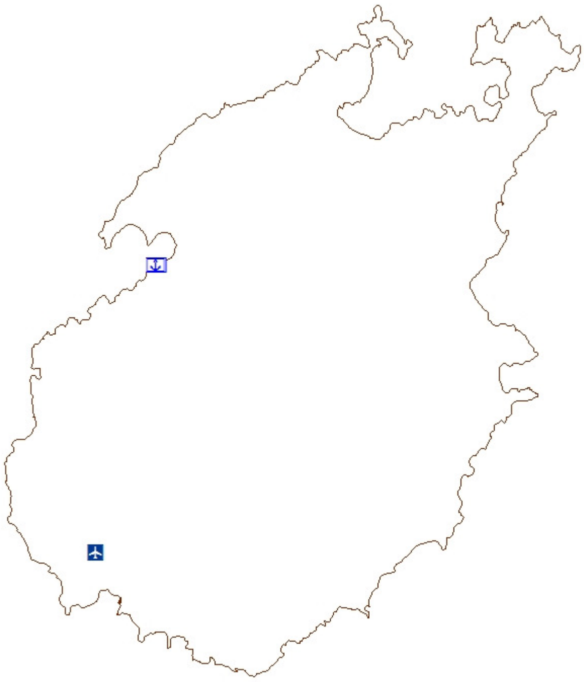

Figure 6.7: Accessibility points of Paros.

6.4 SUITABILITY OF PAROS ISLAND – CONCLUSIONS

In this chapter we stated what Paros island has to offer to its tourists. After getting to know what a tourist mainly wants before starting his trip, we had to check whether Paros fulfils the expectations and preferences they have. For this reason we compared the main needs of the tourists to what Paros can offer to cover these particular needs. In this case if there are decent beaches, some restaurants, archaeological sites and natural parks. As it is already shown previously in this chapter 6, Paros can offer a lot of beaches and many restaurants, which are the two most wanted preferences of the tourists. The next two preferences that are most important to tourists are archaeological sites and natural parks. These sites are present in Paros but are not of the highest importance. There are some archaeological sites which, however, are not of the most important that Greece has to show. In addition, there is one of the few butterfly parks of Greece as well as an environmental park.

In conclusion, we can accept that Paros is an island that meets the major needs of most tourists, making it a suitable case study for this research. So we can now continue with the next chapter of finding which is the best way to estimate specific areas that would be suitable individually for tourists according to the spatial factor of distance.

7. DISTANCE-BASED LOCATION CALCULATION

Up until this point of the research we have narrowed down to a certain group of people whose touristic preferences will be taken into consideration. These preferences were collected in order to find out what data we have to collect for our application. We have also concluded that Paros is a suitable island which fulfils the expectations of the mentioned group for tourists. Since the spatial component of our study is to find which activity locations are closer, we also mentioned the ways that a distance can be calculated. However, by collecting this information and data we cannot take any decisions yet, and that is because this data needs to be processed in order to be able to come up with suggestions about which is the optimal location of Paros for spending summer vacations. In the following sub-chapters we will show the steps taken to build up our application.

7.1 COLLECTING DATA

In order to do some kind of an analysis it is obvious that we need some data to process. According to tourists' preferences, the most important locations for summer vacations are beaches, restaurants, archaeological sites and natural parks (see chapter 5). For this reason, we collected the corresponding data from various sources including hiking maps of Paros, touristic websites regarding Paros locations, the municipality of Paros and the Greek yellow pages.

Since the data collected were all digitized and for the reasons stated in chapter 2, all our datasets are in vector form. We have to mention here that the basic scale on which we work is 1:40.000 because the application is rather general and it does not focus on very detailed information. For this reason there are some point features that only state the presence of a certain place and not an absolute number of them. That is because in a small scale it would be confusing to show every single feature. For example there are many towns where there is more than one hotel. If we had recorded all of them then in our map the points representing them would simply all be almost on top of each other. That is why in every city which includes more than one hotel, we only use one point to represent all of them. The same procedure was followed with restaurants. In the case of beaches, archaeological sites and natural parks every single record of them is stored and used because they are not so many, and therefore they are not getting confused with each other.

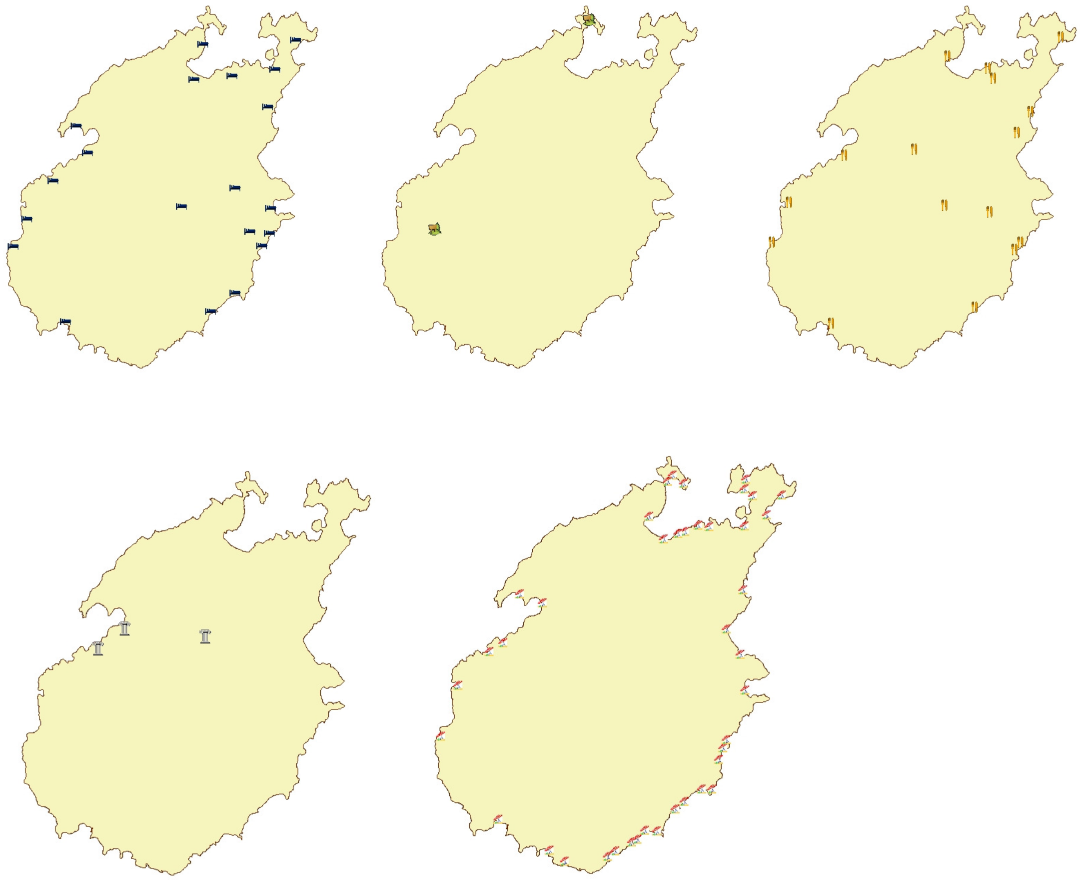

Figure 7.1: Collected datasets. From upper left to lower right, hotels, natural parks, restaurants, archaeological sites and beaches.

The data shown in figure 7.1 are going to be those who will be a major input for the analysis part but more details on how they are going to be used will be stated in the following chapters.

7.2 DISTANCE CALCULATION

The main difference of this application compared to the existing ones is that here the factor of distance is present, which makes it a more spatial application than the others that already exist. Here a tourist can decide which place is most suitable for him or her according to how far away this place is from the areas which are of interest to the tourist, such as beaches etc. In order to make our application crisper, we will assume that the tourist is interested on which area would be best for him to book a hotel. This will strongly depend on what is close to every hotel location. If the locations that the tourist is interested in visiting are within a certain distance from a hotel then this hotel would be a potential place for the tourist to stay during his holidays.

7.2.1 THIESSEN POLYGONS

Our first, and simplest, approach is to apply the Thiessen polygons in order to divide the island into service area polygons which will show which area is closer to every hotel. This way, whatever place is within a certain Thiessen polygon it automatically would mean that this place is closest to the hotel to which the polygon belongs. The result of such estimation would be the one shown in figure 7.2.

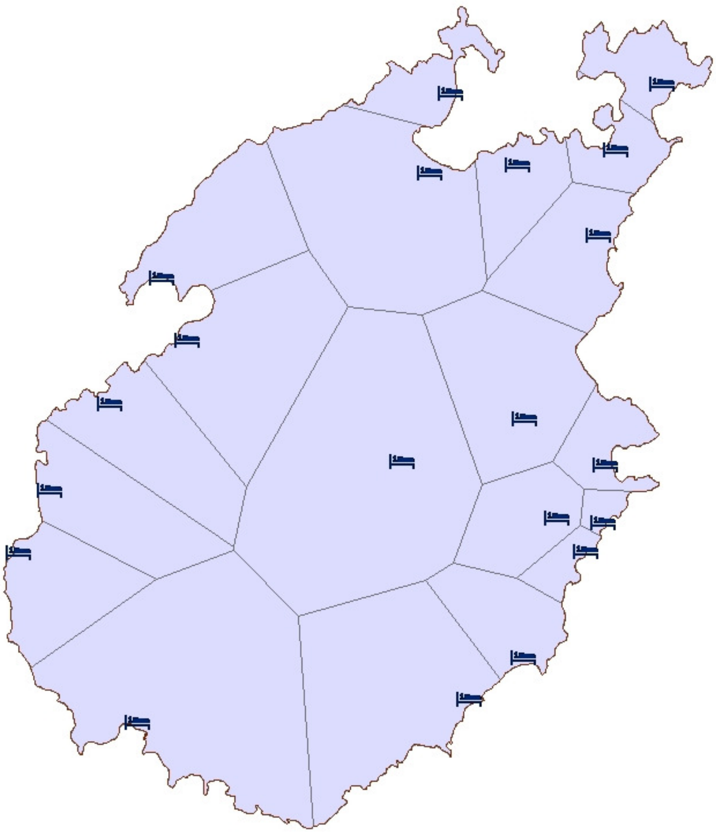

Figure 7.2: Service areas according to Thiessen polygon calculation.

By adding the places of interest to this map we would have the map of figure 7.3 where one can judge more or less what is closer to each hotel. For example, if somebody is mostly interested on going to different beaches then he should probably focus on the southeast or northern part of Paros. If he is not so eager to explore many beaches and he wants to have a combination of beaches, archaeological sites and natural parks then he should rather focus on one of the west side areas of the island and stay to the hotel(s) which are part of those polygons.

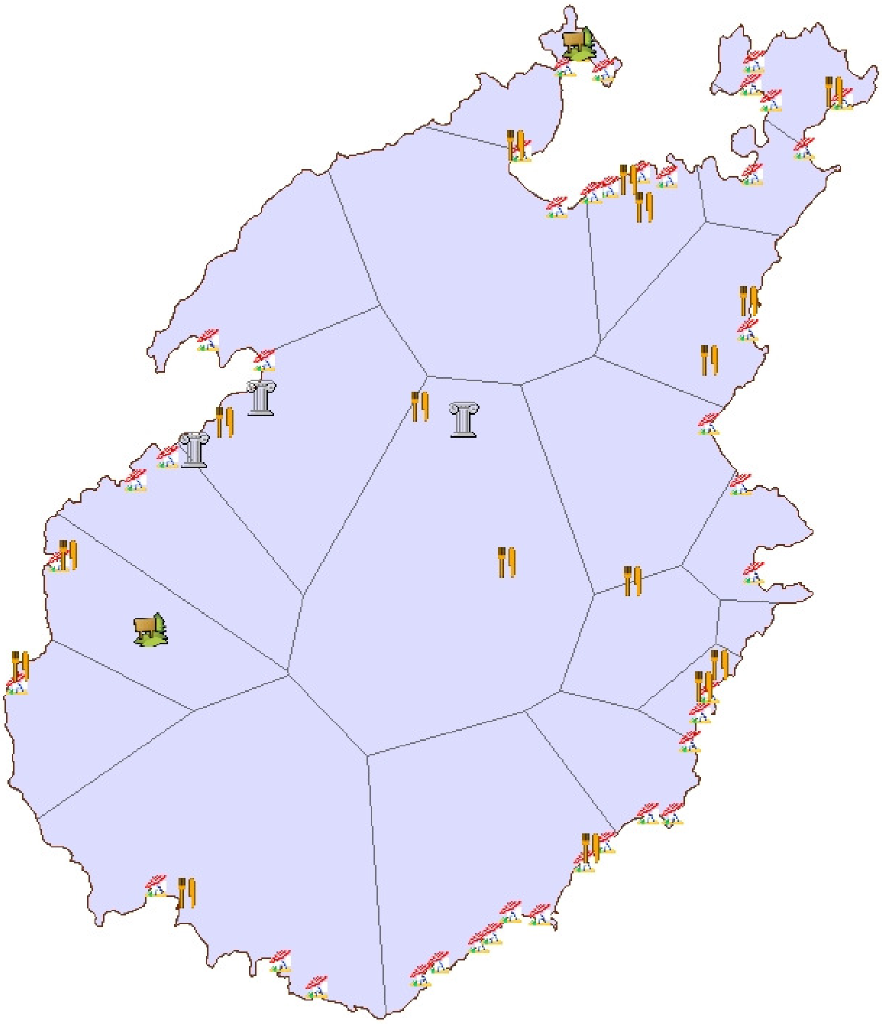

Figure 7.3: Places of interest within the service areas of the hotels.

The generation of Thiessen polygons helps the tourist get a general idea about where he should focus on staying during their vacation. At the same time, these polygons are not the exact reality reflection since the do not take at all into consideration any network which exists on the island such as the road network. Another problematic situation that has occurred is at the northeast area of the island as shown in figure 7.4.

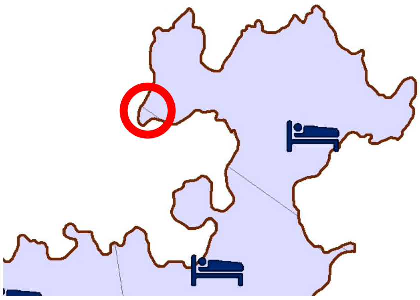

Figure 7.4: Problematic situation.

Even though we have mentioned that every polygon shows the closest areas to the hotel which is included in this polygon, in this case it is obvious that the small circled area should be part of the northern polygon because one should anyway go around the gulf in order to reach it. This error happens because the Thiessen polygons only calculate straight distances without taking into consideration any borders which in this case are the coastline. If there were more hotels in the northern part of Paros there would be more problems with such kind of errors. A solution to this problem would be to use another way of calculating polygons like for instance based on a road network.

7.2.2 BUFFER ZONES

Our second approach is to use buffer zones. In this case a buffer will be created around every point of interest, according to how far away the tourist wants to be from it. These buffers will be combined in such a way that in the end we will have a result showing which areas are most suitable, the ones which are not suitable and of course those which are somewhere in between. In the following chapters this method will be explained in detail.

Let's assume that a tourist prefers to stay at a place which is 20Km away from a natural park, 10Km away from any archaeological site, 700m away from any restaurant and 300m from a beach. In this case, our main goal would be to create buffer zones around each layer's features according to the distances specified respectively. After calculating these buffers, new polygons will be created around each feature, which polygons might intersect with each other. We have to state now that we will deal with straight line buffers, which will result in perfect circular buffer zones. If the buffers of two or more different layers intersect, then the common polygon is the area which fulfils the criterion of being within the specified distance of the places whose buffers intersect. For example, if a buffer polygon around a beach (300m radius) and a buffer polygon around a restaurant (700 m radius) intersect, then the common polygon represents an area which is both maximum 300m away from the beach and maximum 700m away from the restaurant. Accordingly, if more than two buffers intersect, then the polygons which are part of every buffer are the areas which fulfil all the criteria. The less buffers intersect with each other, the less criteria are fulfilled. Figure 7.5 uses a graphical way to explain this concept and makes things a bit clearer.

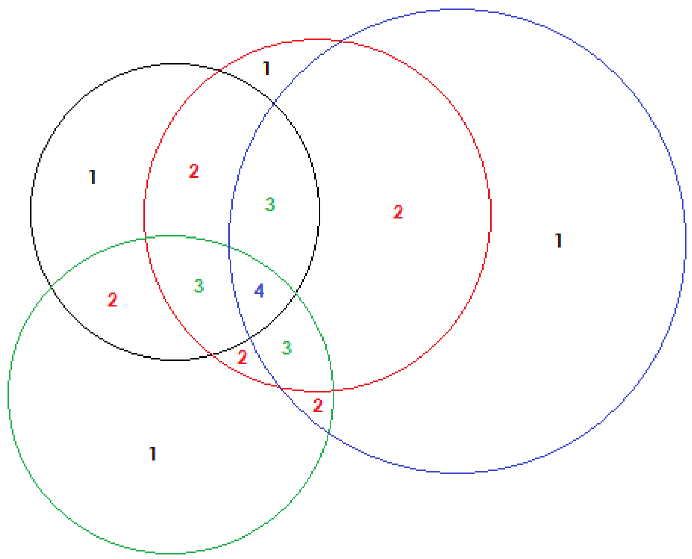

Figure 7.5: Intersection example.

In figure 7.5 we see possible intersections between 4 polygons which could as well represent buffers from four different features from different layers, like buffers around a beach, a restaurant, an archaeological site and a natural park. In this figure we can see the number of layers that intersect in different areas. The numbers in the formed polygons show the count of buffer zones that intersect at the particular area. The result of these intersections could be represented in a bit better way (figure 7.6) by giving colours to the areas according to the number of intersected buffers.

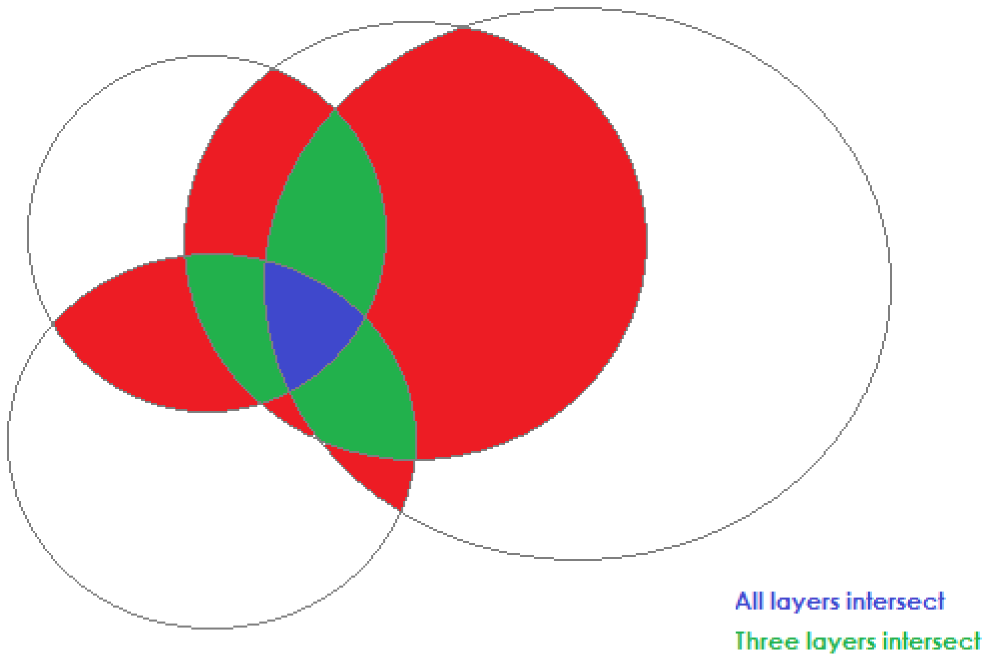

Figure 7.6: Distance analysis result.

In figure 7.6 we can see how the result of the distance analysis could look like, according to the fulfilment of the criteria. The blue area would be the area which fulfils all the four distance preferences for the four layers; the green area fulfils three of them and the red only two. So, if this represented real parts of an area, the tourist could automatically see which area is best or worse for him.

The above explained method will also be applied for our dynamic application. In this paper by saying dynamic we mean that the user can interact with the application, by setting his preferred distance values to certain locations, and that it does not just give a default result for everybody. In this case, the user would fill in the maximum distance that he/ accepts of being far away from beaches, restaurants, archaeological sites and natural parks. After doing so, he will receive a map result of Paros island and the areas which are from most to least suitable for his personal preferences. This is going to be achieved by building a model, which will be specific for this multi-criteria analysis and which is going to be explained in more detail in the following pages.

7.3 DISTANCE ANALYSIS APPLICATION

In this chapter we will show and explain the application, how it was built and how it works by having a look to almost all the steps taken to achieve a final result.

Our goal, as mentioned before, is to provide tourists with a solution regarding where they should go for summer holidays within the total area of Paros island, according to their individual preferences. And by preferences we especially mean the spatial preferences of distance to certain points of interest. The user has to specify how far away he wants his hotel to be from certain locations, e.g. beaches and restaurants. With our application he will be able to decide which area of the island would be best for him/ her to book a hotel, which areas would be more or less ok and which would not be suitable at all. The result would be something like what is shown in figure 7.6. But in order to reach this stage a whole model has to be run which includes some steps which will be described below.

In order to have a general overview of what is going to be done within our application we first have to see the complete model as shown in figure 7.7.

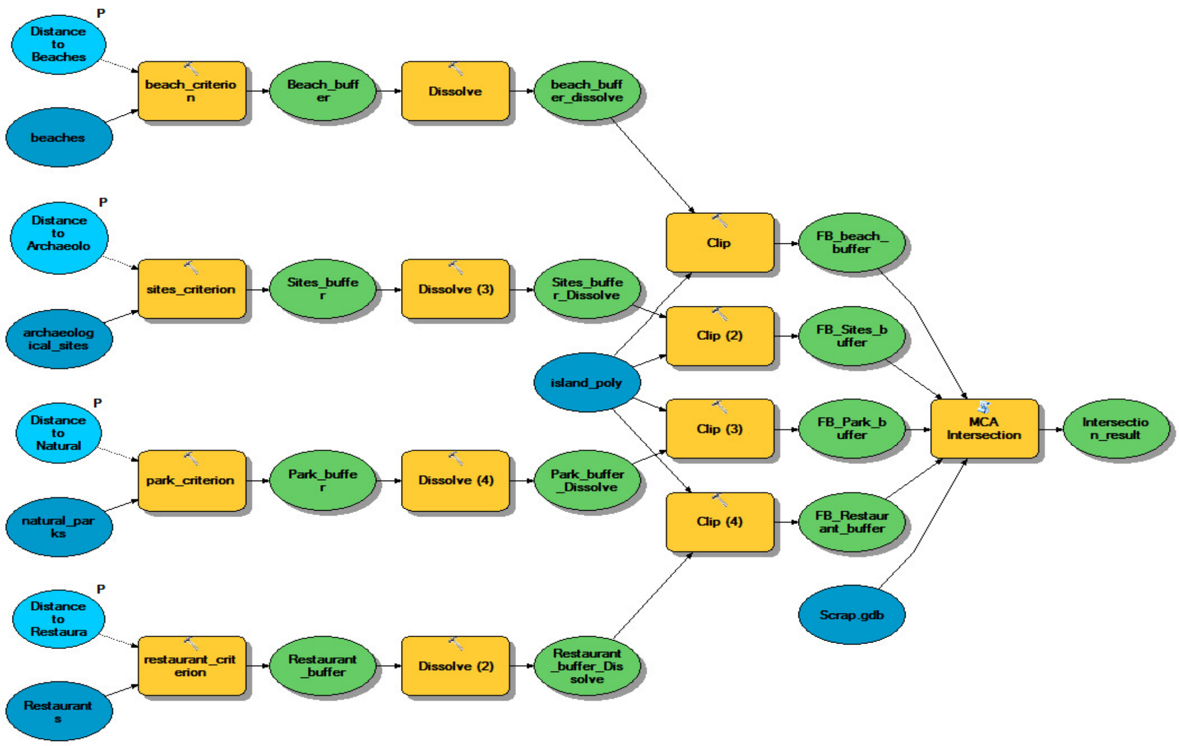

Figure 7.7: Implementation model.

As one might notice on first site, most of the model consists of simple, already existing in ArcGIS, tools. In a few words, what this model does is to get the specified preferred distances from certain locations, and according to these specifications it calculates the final result by calculating the most suitable areas for staying according to the input distances. But let's see now how this would work step by step.

As explained previously, our first goal is to create buffer zones around our existing points of interest, according to the individual tourist's input. Therefore, we create buffers around every layer according to a certain distance value set from the user. We have to mention here that the buffer distance value is set to be an external variable so that the users can fill in their personal preferences.

At the same time, we have set a precondition to the buffer tool which makes sure that a certain buffer will only be created if the user puts a value greater than 0. We included this detail because there can be cases where a user does not really mind how far he is from a certain location.

If this precondition was not set then our buffer tool would try to create a 0 buffer zone which is not possible and would cause our model to crash, give an error message and terminate the whole procedure. Of course we do not want that because one of the suggested layers might not be of interest to a certain tourist but the buffer zones around some other layers will have to be calculated. So, in a few words, by setting a precondition we manage to skip the buffer tools which have 0 as an input value and continue with the rest of the model.

Since we have created all the buffers around every feature layer of interest, theoretically we could start intersecting them to find out the places where most of them intersect. However there are some intermediate steps that need to be taken in order to have clearer results as well as to minimize the volume of calculations. After creating buffers, it is very likely that we have polygons lying on top of each other which results in having duplicate information for the same area. For this reason we first dissolve all the buffer layers separately so that we only have one record for each one of them and therefore no duplicates. An example is shown here in figure 7.8, where the buffers around the beach sites are dissolved into one feature.

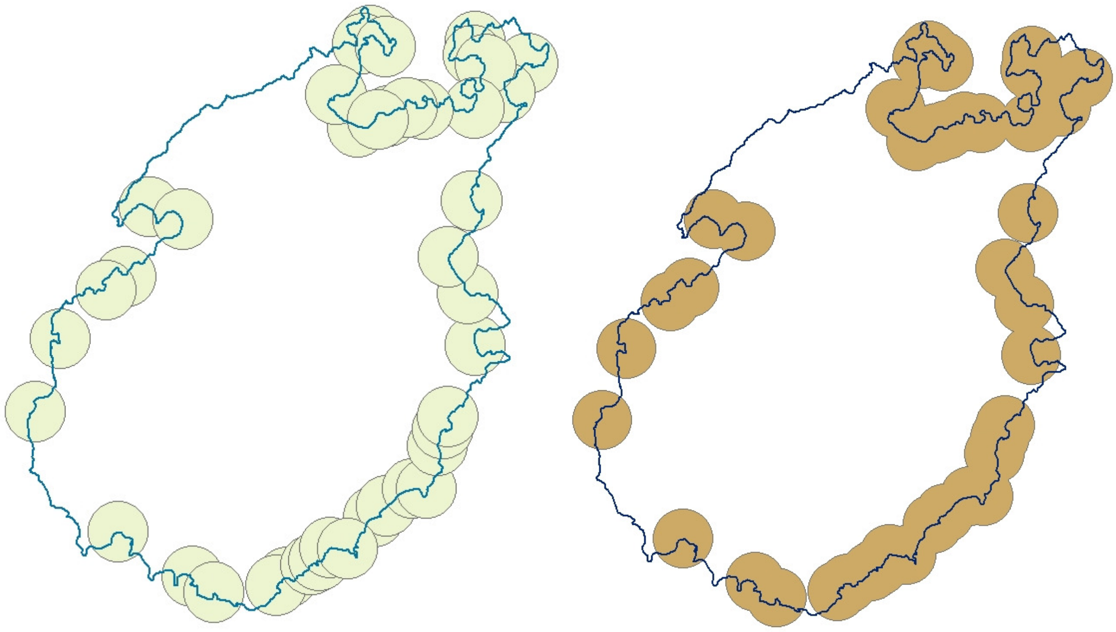

Figure 7.8: Buffer layers before (left) and after (right) dissolving.

We have already minimized the size of or data by dissolving it but as one can see, there are still parts of the (dissolved) buffers which are useless and these are the areas which lie outside of the island's borders. If we leave them as they are, they are going to be part of the following calculations even though nobody is ever going to find a hotel in the sea and this will slow down our model for no reason at all. That is why we will also clip the buffers according to the island's shape. The output of each layer's clip is finally going to be the layers that are going to be intersected in order to calculate the best vacation locations (see figure 7.9).

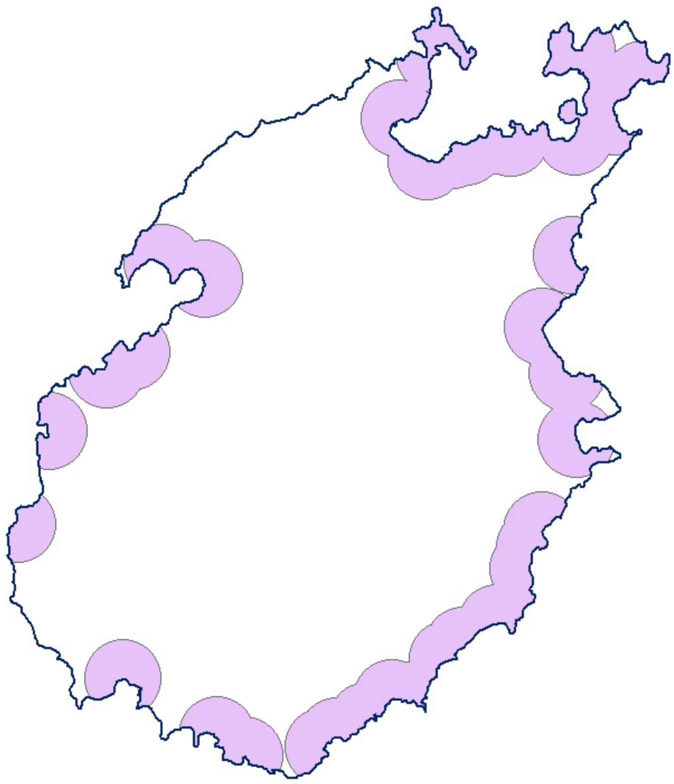

Figure 7.9: Potential final buffer layer for beaches.

The exact same procedure will be followed for all four layers, given that the user has indicated a distance to every single one of them. If not, then the features that the tourist is not interested in, will not take part into any of the calculations.

Until now we explained some simple preparations that needed to take place so that we optimize the volume of our application's input and minimize the time which our model will calculate the results. The final buffer layers will be the input of the most important part of the model, the core of it, where the optimal locations are actually calculated. Given the complexity of vector calculations, it is impossible to calculate the optimal areas the way it is shown in figure 7.6 only by using model builder's abilities. It is a complicated calculation which even if it would be possible to calculate with model builder's tools the model would be extremely complex and hard to understand. For this reason, we decided to create a script which will do the same job with the big difference that it is much more flexible and it supports easy ways of iterating, customizing and controlling the whole procedure. We are now going to explain the basic steps of this script, which is a very important piece or our model and which is shown in appendix 5.

In general, our Intersection script calculates in which areas of the island Paros there are intersections of the distance buffers explained above. Every area is assigned with a certain 'level' number according to how many criteria are fulfilled (meaning how many distance buffer areas intersect at the particular area). Here we have to mention that already the initial point layers, containing the places of interest, are supplied with a field called 'level'. In doing so, every layer in the end will also have a 'level' attribute which will give us information, in our final result, about the number of intersected areas, and therefore the suitability level of that area. But this will be explained more in detail at the explanation of the corresponding part of the code.

Getting a little bit more into detail, the goal of our script is to intersect every layer with any other layer[3], including those layers which contain polygons generated from intersections.

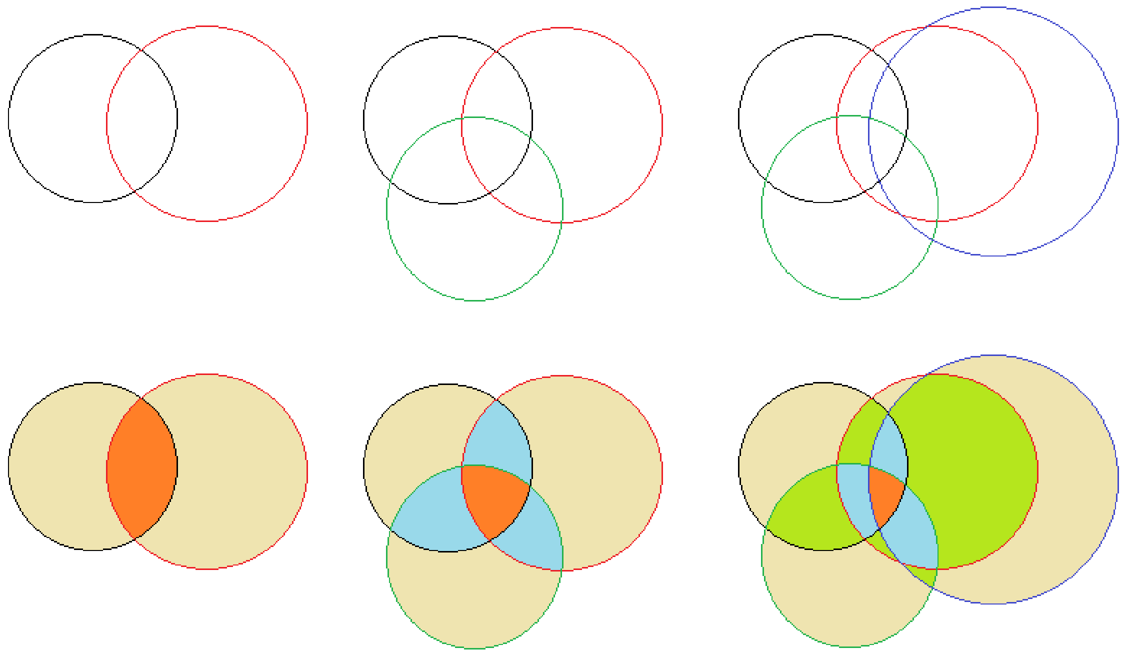

Figure 7.10: Steps of intersecting layers.

Figure 7.10 shows more clearly the steps that need to be taken in the case that we have four layers in total, which are supposed to be the buffer layers around our four places of interest: the beaches the restaurants, the archaeological sites and the natural parks . So first of all, we have to intersect the first two layers, let's say the beach and the restaurant buffers. Since initially our layers are assigned with a level

[3] From this point on, when we lay layer we refer to the optimized buffer polygons created at the first stage of our model, before entering the MCA Intersection script.

value of 1, because there is no intersection yet, their common (intersected) area will get a level value of 2 because two layers are present at that area.

In the second step, a third layer, let's say the archaeological sites buffer, comes on top of our previous layers (the two initial layers, and the intersection part). We now have to find out where there is an intersection between the archaeological sites buffer and the rest. In practice, in this step there are three different intersections: between the archaeological sites buffer and the beaches buffer, between the archaeological sites buffer and the restaurants buffer and finally also between the archaeological sites buffer and the intersection between the beach buffer and restaurant buffer. At the same time, the level values of the intersecting layers will be added up so that the new intersected area results have a higher level value, since there is one more layer present. At the end of step two, and according to figure 7.10, there is one area where all three layers are present, which is the orange coloured part. In addition there are three areas where only two of the layers intersect and which have the level value 2. These are the blue coloured ones. Finally, there are some areas where only one buffer layer is present, the light brown ones, and which only have a level value of 1 since there is no intersection at this area.

Now the third step, where one more layer (let's say the natural parks buffer) comes to overlay to the rest, makes things more complex. The concept stays the same, that this new layer has to get intersected with all the single layers as well as with all the previous intersections. Since every new layer has a level value of 1, when adding this value up to the level values of the rest of the layers we will get a new hierarchy in our result which is shown in the right part of figure 7.10. The orange part is the area where all four layers intersect, and therefore has a level value of 4, the blue parts represent the areas with level 3, which means that three of the layers intersect, the green parts are the areas where only two layers intersect and therefore they all have a level value of 2 and finally the light brown ones where only one layer is present.

Now all this explanation has to be translated into code in order to actually be able to do these calculations in our model. We had to find a solution where every new layer would have to intersect with all the previous layers and also with all the previous intersections in order to update the situation of the total area. To do so we came up with the solution of working with two lists. One of them will contain all the initial layers and the other one all the intersection results including also the initial layers. The first one we call it LayerList and the second one ResultList. At the beginning, the two lists will look like this:

LayerList[A,B,C,D][4]
ResultList[ ]

The LayerList will be ready, filled with all the initial buffer layers, whereas the ResultList will be empty at the beginning because there is no result yet.

The basic part of the code intersects layers and fills in the ResultList with intersection results and the initial layers and can be roughly shown with a small example code part which is simplified for easy-understanding purposes[5]. This example code looks like the following lines:

```
for Layer in Layerlist:
      append (Layer, Layerlist)
      for ResultLayer in Result Layer:
      if Layer != ResultLayer:
            Intersection = Intersect (Layer, ResultLayer)
            Append (Intersection, ResultLayer)
```

In this example piece of code, there are two loops. The first one goes over every single record (Lalyer) of LayerList and appends the layer to the ResultList. This makes sure that all the initial layers get added to the ResultList. The layers have to be added there because every initial layer has to be intersected also with the rest of the initial layers. Now there is another loop, in the first loop, which goes through the ResultList which actually makes it possible to combine every layer form LayerList with every layer from the ResultList. In this loop the intersection takes place as well as the appending of the intersection result to the ResultList. We have to note here that there is a condition in this loop that states that the intersection will only take place if the two intersecting layers are not the same layer. It would make no sense to intersect

[4] Capital letters replace the layer names here for simplicity reasons.
[5] This example code is only for understanding purposes and therefore the syntax of it is not executable by python.

layer A with layer A. In the following lines we will show how this iteration works by showing the state of every list by the end of every loop of the LayerList.

Iteration 1

LayerList[A,B,C,D]

ResultList[A] *no intersection will take place*

Iteration 2

LayerList[A,B,C,D]

ResultList[A,B,AB] *there is an intersection between A and B*

Iteration 3

LayerList[A,B,C,D]

ResultList[A, B, AB, C, AC, BC, ABC] *there is an intersection between C and all the contents of ResultLayer from iteration 2*

Iteration 4

LayerList[A,B,C,D]

ResultList[A, B, AB, C, AC, BC, ABC, D, AD, BD, ABD, CD, ACD, BCD, ABCD]

There is an intersection between D and all the contents of ResultLayer from iteration 3

One can now see it more clearly that these two loops help us manage to intersect every layer with all the rest layers, initial ones and the existing intersection results. All the layers within the ResultList in a combination will give us the final result that we are looking for. But at this point we will merge all these files into one layer for simple management and display reasons. This layer will contain all the layers or the ResultList as different records. We will finally colour all these records according to their level value and we will have our final result in its final form, ready for the users to see. By adding to the result also the hotel locations, the user will be able to see which of those hotels are closest to some or all of their places of interest. A potential result for a user could be an output like the one shown in figure 7.11.

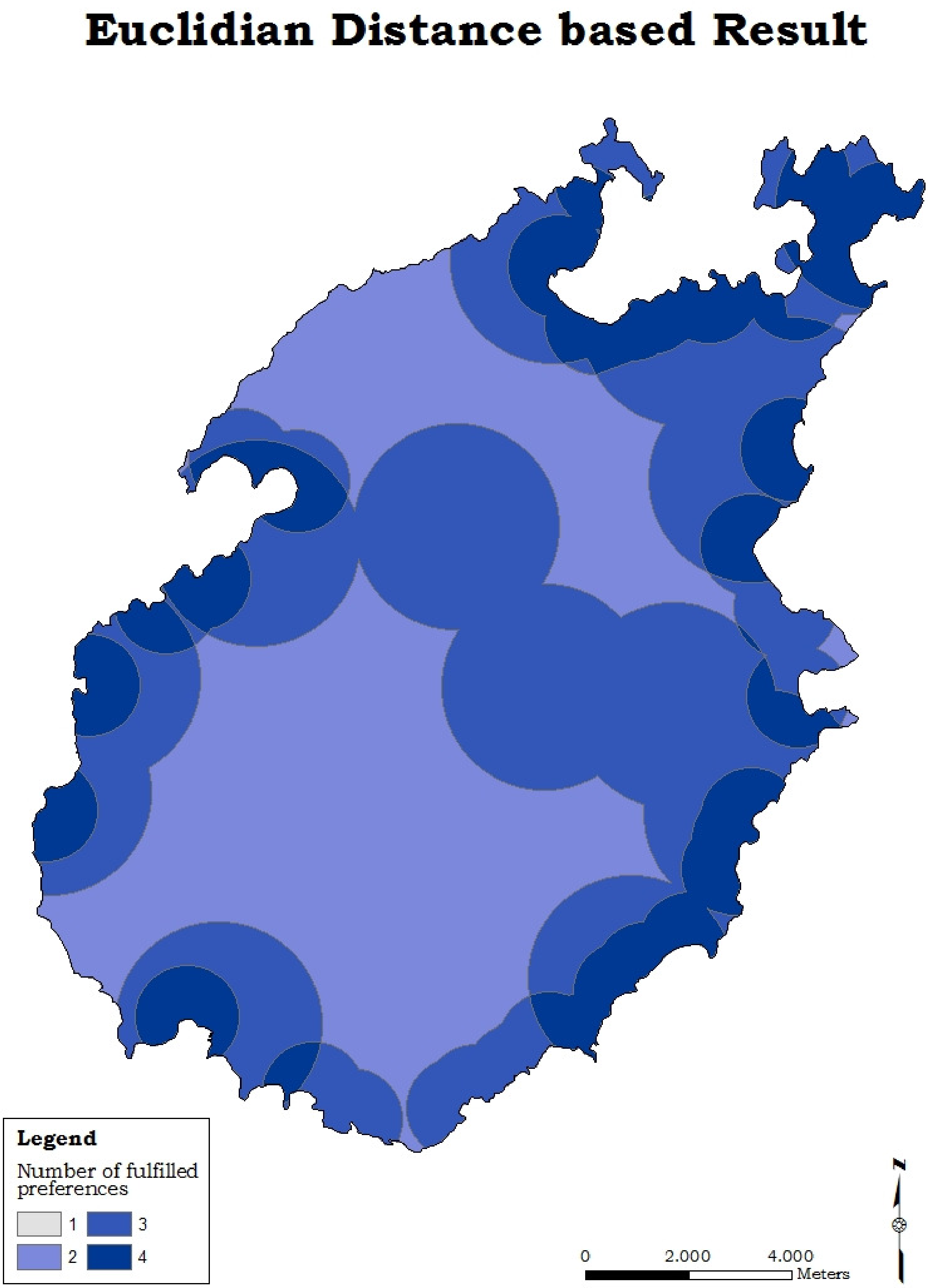

Figure 7.11: A potential Euclidian Distance based result.

In the example output of the distance based analysis, as shown in figure 7.11, the darkest blue areas are the ones which fulfilled all four preferences of the user, the next darker blue ones three of them whereas the light blue ones only two. By showing also the hotel locations the tourist user can decide which hotel would be the best choice according to his/ her preferences. Even if the existing database of hotel locations is not up to date at some point, the user can still see the areas where he should generally look for one.

7.4 ADDITIONAL DETAILS

Up until now we have explained what the main part of our code does in general. However, there are some more details that make our script more flexible and error-free. In addition, there are some parts which make sure that we get rid of unnecessary parts of our layers.

To start with, we have to mention that after executing an intersection, except for updating the level value of the layers, we also dissolve the result so that the result of the intersection will only have one record which will include all the different parts of the intersection. This will make our script faster when it will have to include this layer to another intersection.

A very important part which is included in the script and plays a very important role in getting rid of unnecessary data, is the part where duplicate records get deleted. For example, after every intersection, the result of the function will be deleted from the layers involved in the intersection. Looking at the first step in figure 7.10, the orange part will be deleted both from layer A and B. That is because this area is already a new area called AB (from the intersection of A and B). There is no reason anymore for this area to be part of the two intersecting areas because it is anyways going to be overlapped from the result of the intersection. This procedure is done in every intersecting pair in the script. The intersection result's area will always get deleted from the two layers that were part of the intersection. In this way the layers will be smaller the next time that they get involved into a function. The following illustration (figure 7.12) makes this concept a bit clearer.

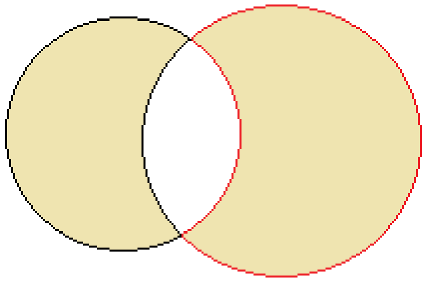

Figure 7.12: Erasing the intersected part from the involved layers.

As we can see, the part which is common in these two layers is deleted from both of them and in the next intersections that these two layers will be involved they will just be in this shape. If a layer wants to intersect with the blank space that now exists between these two layers it will be intersected with the intersection result of them which is now filling this gap. The same concept will be applied in all the intersections that will take place.

Another important issue is that during this process a lot of intermediate results are going to be generated. If these results would get stored to our normal database where all the rest of the data is stored then the database would get full of unnecessary files which are only used for calculating the final result and not for any other reason. That is why we decided to store these files in the memory which will let our basic database free and also our model will be faster since data is not going to be written on the hard disc of the computer. However, the procedure of renaming the intersecting layers is not possible to be done because this function is not executable using the memory space. For this reason, an additional database called Scrap.gdb is created. There we are going to store all the intermediate files which cannot be stored in the memory space.

Last but not least, having in mind that a user might not care about one or more layers, the code was adjusted in such a way so that if there is zero input for one layer, it will just ignore this layer and proceed with the rest, just like this layer does not exist and will not be part of any calculations.

7.5 NETWORK-BASED SOLUTION

Our initial analysis was based on calculating straight line distances without taking into consideration any roads. However, in real world it is almost impossible to go from one place to another following a straight and direct line, especially in built up areas. For this reason we will also estimate optimal vacation areas based on network distances. In this case, the buffers created around the points of interest will be calculated according to the road network which lies around them. In principal, the whole method will stay the same with the only difference that the buffers will not be created based on straight line distances but in network distances (see Figure 7.13). The result of this analysis is shown in figure 7.14.

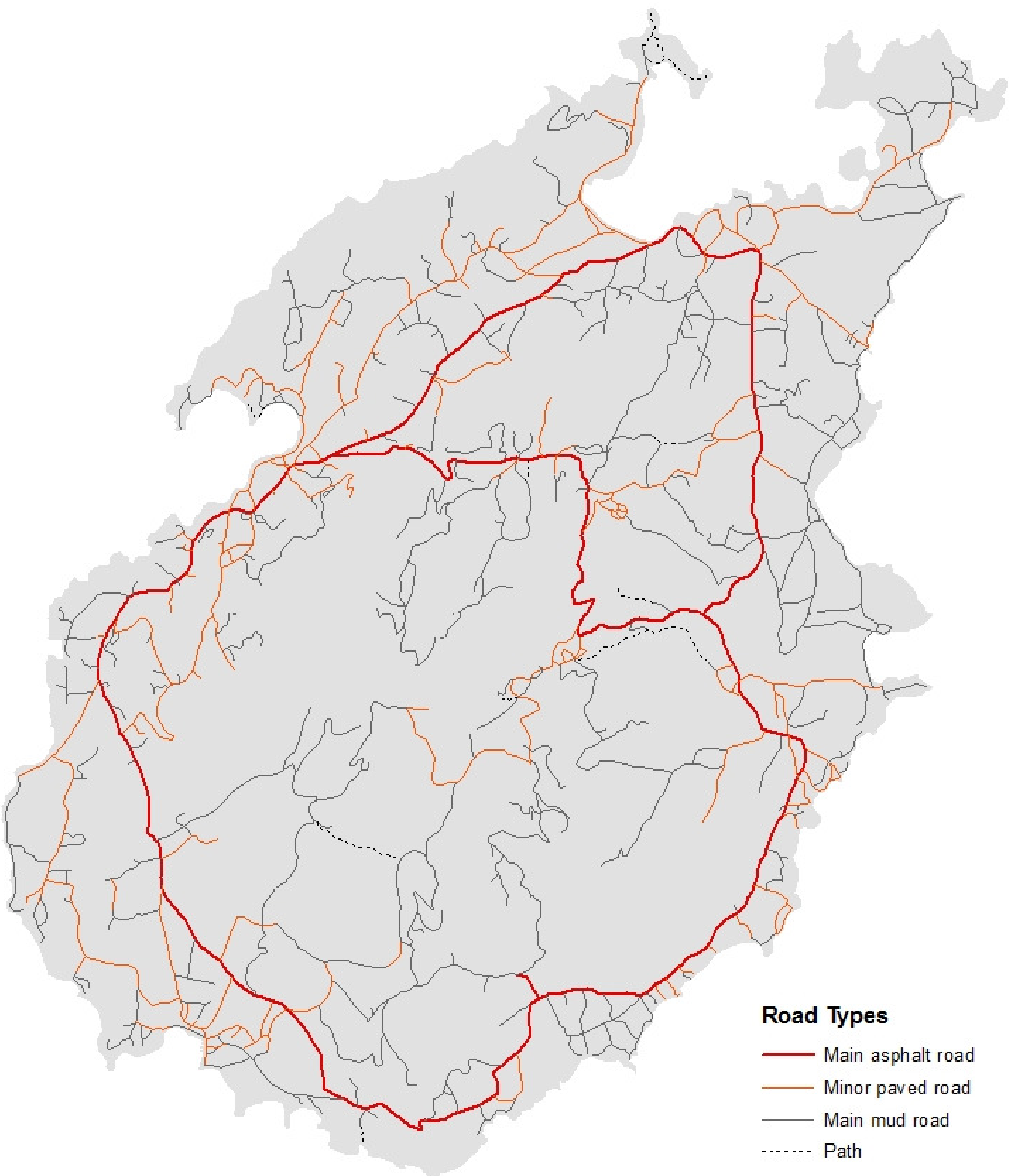

Figure 7.13: The road network of Paros.

Straight distance calculations could lead into false results in certain special cases. There can be solutions to this problem. The easiest one is to calculate suitable areas based on distance calculations using a certain network. For instance, using the road network as well as pedestrian paths. Buffer zones will then be replaced with polygons around the points of interest which will represent the service area of these points based on the network that is around them. This will minimize the errors created by straight distance calculations and will be more realistic. In that case we would have to make sure that there is a sufficient network connection all the way to the points of interest, otherwise we will be forced to use a combination of straight line and network distance calculations.

We have to mention here that the preferred distances to the places of interest are the same so that we can compare the differences in the two results. However, it has to be mentioned here that since we only focus on physical distances, all road types are considered to be equal. No matter if it is an asphalt road or a pedestrian path it doesn't matter because we only care about distances through the network. As we can see from figure 7.14, the results are somehow altered due to the fact that in this case one has to follow a certain network to go towards a certain place. Since it is understandable that moving though a certain network makes distances bigger, it is natural to expect that the buffers are smaller for network distance calculations compared to those calculated based in Euclidian distance when the same distance is specified.

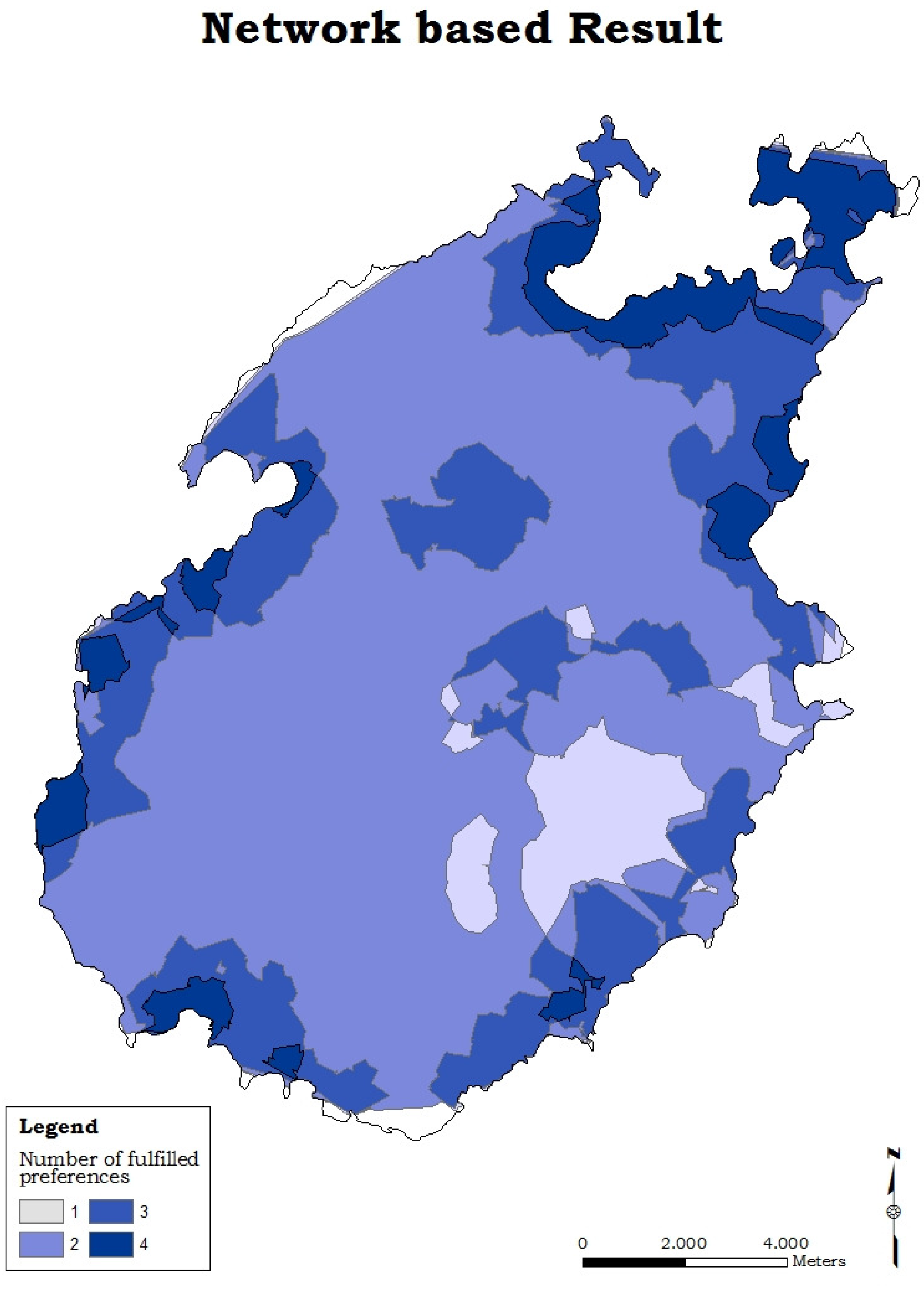

Figure 7.14: A potential Network Distance based Result.

We can notice that even though we used exactly the same distances to places, in this case there are some areas of the island which are not suitable at all for the given preferences, whereas in the straight distance solution there was no such an area. And this is again a result of having smaller buffers around the places of interest.

7.6 CONCLUSIONS

After having created and used our application we can now say that there is a new way of helping out tourists to find their personal optimal location for summer vacations in Paros island. This application provides a very exact result based on the user's input which is also graphically shown in an easy to understand map. No matter what the input is, the model will generate a suitability map, except of course if a user does not fill anything in.

The two different approaches give us results of different detail as well. The straight distance calculations give a more rough result whereas the network based one offers a more detailed result. If we take a look at the results by putting them next to each other (see Figure 7.15) we can see some clear differences.

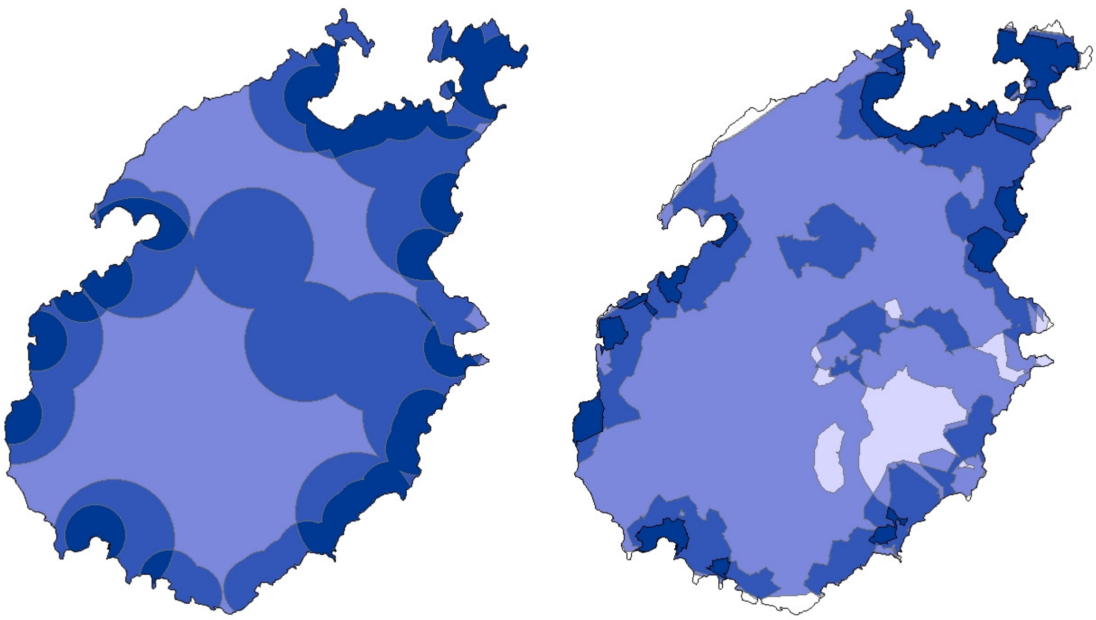

Figure 7.15: Results based on Euclidian (left) and Network (right) distances.

The Network based result is closer to reality but that does not mean that it can also fully represent reality of course. As we can see, using the road network to calculate distances has as unavoidable result smaller areas. If we look at the southeast side of the island there are almost no areas that fulfil all the criteria anymore and that is obviously because the buffers are now smaller and therefore in some areas they do not intersect anymore. Especially the southeast side of the island has very different results. Even though in the first approach (straight line distance) there was a wide area where all 4 of the criteria were fulfilled, when using the network-based approach less criteria get fulfilled. That is because the service areas of the places of interest are smaller and do not reach that area. An example is the service area of the natural parks. As a conclusion, we can prove now that using the same preferences, the suitable areas calculated by using the road network are smaller. In order to have a more quantitative comparison of our results we will calculate the area values of the different level regions. Table 7.1 shows in more detail these area calculations.

	Level 1	Level 2	Level 3	Level 4
Buffer Result	-	80.35 Km^2	79.80 Km^2	36.55 Km^2
Network Result	12.31 Km^2	118.85 Km^2	47.05 Km^2	16.08 Km^2

Table 7.1: Area values of the results.

As we can see from the above table, the two different approaches have big differences also from a quantitative point of view. Our first comment is that in the buffer result there are no areas which fulfil only one preference, whereas in the network based result there are 12.31 Km^2 which are within the distance limits to one place of interest. In the same way, the areas which fulfil two distance preferences are again bigger in the network-based result than in the buffer-based one. On the other hand, the areas which fulfil three to four distance preferences are bigger in the buffer-based results than the network based ones. We have to also mention here that in the network-based result there are also areas where no preferences are fulfilled whereas in the buffer-based one this is not the case. In general, in the network-based result the areas are smaller because they are calculated based on the road network which is of course not in straight lines.

Even though the network based result is much more realistic, that does not mean that it fully represents reality as other factors can still influence distances, such as economic, temperature, terrain factors. A very basic factor is the terrain which also influences the distances. It would definitely take more time to go from one point to another if there is a big hill in between than when the area is completely flat. If the terrain would have also been included in the calculations, then the buffers would appear even smaller and so the final areas would be more limited. The advantage of this application is that once the buffers are created, the calculations will remain the same since it deals with any kinds of polygons. For example if one creates service areas based on a road network also taking into consideration the terrain. The result in that case would be different but the good thing is that the calculations would stay exactly the same and so our code would not have to be altered.

We also have to mention here that one should not set a very low distance value for calculating suitable areas and that is because that will make the results very inaccurate. Since our initial collected data is based on a 1:40.000 scaled map, the location of each placed point cannot be very accurately placed because the scale is rather small. It would be best if the minimum accepted value was around 100 meters. Then we are going to have useful results. If this number is smaller then there is a very big chance that there will be no areas that fulfil more than one preference.

8. CONCLUSIONS, DISCUSSION AND RECOMMENDATIONS

After going through all the necessary steps to give a solution to our problem, which was to calculate suitable summer vacation locations for Paros island, we came up with some useful conclusions extracted from every research part of it.

8.1 TEST GROUP AND PREFERENCES

First of all, forming a test group was essential for setting a crisp and solid base of our research. This helped us in a very big extent to be more precise about what we are going to work on. In a research such as a thesis work it is very helpful to be able to focus on a certain part because otherwise it can get complicated and much broader than expected. And this is exactly the reason why forming a test group is of a great help. In our case this group included people who are in couples, have no children, are EU or Greek citizens and are aged between 25 and 64 years old. If this step would have not been done then we would have much more diverse preferences since every individual has different interests. If one has to be very precise though, he could argue that even within a certain group of people, opinions can vary but this is something that we did not take into consideration.

Gathering information about the tourists' preferences helped us focus on a few criteria that would be the input of our analysis. More specifically, we came up with the conclusion that most tourists want to have beaches, restaurants, archaeological sites and natural parks around them when they are on summer vacation. Focusing only on these four different kinds of places of interest, made it less complex to calculate the final optimal areas for vacation. If we had used every single preference of every candidate, our model would be much more complicated and we do not want that because in this study we only try to find out how this whole application could work and not to make it complex. Besides, in our case additional preferences, which would automatically mean additional layers to be included in the analysis, would not influence the calculation flow much. However, the more criteria combined, the more the result will fit to the user's complete preference list but this is something that could be implemented in a future version of this application.

8.2 METHODS AND DISTANCE BASED APPLICATION

First of all we have to mention here and comment on our choice of choosing to use vector format instead of raster. It is clear that raster is much easier to make calculations on by using raster algebra. However, in our case it was not the optimal solution. That is because our initial obtained data scale was rather small, meaning 1:40.000 which means that there is already some inaccuracy of data. If we used raster this accuracy would increase, especially if the pixel size was rather big, in which case all the results would not be accurate. Vector is undoubtedly more complicated in calculations but much more precise in the results. So if out input data are not of very high accuracy, at least we should use a method that will not deteriorate the situation like raster analysis would. In addition, in our case it is not so complicated since the maximum amount of layers that will be used in the calculations is only four. In case these layers were much more than we would have to adjust the input data, mostly to find much more detailed data sources, so that we can then use raster for the analysis.

As we already mentioned in chapter 3, there are many ways different 'kinds' of distance, including straight distance, distance over a network, time distance etc. In our approach we used straight distance as a basic start. The two alternative options that we used for calculating distances were creating Thiessen polygons as well as generating buffer zones.

The first solution, on one hand, could give some results on what is closer to each hotel so that one can get a rough idea of what he can do around every hotel in Paros. The advantage of this method is that it is very simple in calculations and the user does not have to think about distances himself. On the other hand, Thiessen polygons also have some downsides. First of all, as we saw already in figure 7.4 there can be some miscalculations (that in this case cannot be avoided) due to the fact that the island has borders and these borders can form gulfs and peninsulas. These miscalculations can lead into false conclusions because everything is based in straight distances, without taking into consideration the shape of the island. In addition, the whole analysis that is a major part of this thesis cannot be included if we use Thiessen polygons. The tourists do not have the chance to specify what is important for them to have around the hotel that they will stay and even more they cannot set their own distance preferences to their places of interest. In general, Thiessen polygons can only be used as a very rough calculation.

Our second approach was to use buffer zones around places of interest that are specified by the user according to how close he/ she wants to be to a certain place. These buffer zones are the criteria for each person about the optimal location for vacation based on how far away certain places of interest are. The buffers were then intersected in order to provide the user with a result that indicates which areas fulfil all of the criteria, which fulfil some of them and which none at all. This approach is very useful and very effective since the results are calculated according to every individual's input. It can give customized results and can be applied to any user's preferences. In addition, with the right colouring of the results, the user can really see where is the best area for him to book a hotel and which is the least interesting for him or her. In a few words the advantage of this approach is that it can really be user oriented and very precise regarding the results. However, there are a few points that need to be taken into consideration. The first one is that in our case the buffers are created based on straight distances and network distances which are solutions but they definitely do not fully represent reality because there are other factors which still influence distances such as the terrain of an area. At the same time, one of the problems that was in the Thiessen polygons approach also remains a problem in this case. This is that the distance calculations, through buffers this time, are still not taking into account the shape specificity of the research area. For this reason straight distances also include going through the water. This will lead again to miscalculations in areas that are shown as close to a certain point but in reality it is impossible to go there in a straight line and therefore, the area is not close to the certain point anymore.

The third and last solution to our problem was to calculate areas that fulfil the preferences via the existing road network of the island. By solving the problem with this method we saw that the result was very different. Some areas fulfilled completely different number of criteria that the previous method and some areas did not even fulfil one criterion. By following this procedure we can see that using the same distances in two different approaches can give us completely different results. In our case the last result is the closest to reality. If one would base their decision on the straight line based suggestion it would be more or less acceptable but definitely not as precise as the network based one. Now if some more details would get added into the calculations such as the terrain on which the network is, we would have different results again. But this is something that was not taken into consideration within this study.

8.3 DISCUSSION AND RECOMENDATIONS

In this chapter we will discuss in a few words some issues that were raised during the process of completing this thesis research. These mainly include alternatives for tackling the problem as well as future recommendations to improve the created application in a way that was not possible within the time limits and extents of this study.

To start with, this research only had as a target the island of Paros. But in reality, especially because Greece is in general a touristic country, this application would be useful if it could take into consideration the whole country on maybe even the whole world. Understandably, this would exceed the limits of a master thesis but it could be an independent project for supporting the Hellenic Tourism Organisation as well as the local touristic areas of the whole country of Greece. At this point we can bring up the issue of vector and raster data formats. If raster was our choice it would on one hand be a smart decision from a calculation complexity point of view but at the same time it would be a lot of unnecessary data since not every single part of the country's surface would contribute to the calculations. Setting the resolution of the grid would be a complicated issue to be solved. Or would it be more preferable to use vector? This is a very important question that would have to be answered at an early stage of such a project.

Another point for future optimization of the current application would be to make it available online. It could be a thin client online application since we cannot expect of course that the users would have the appropriate software installed in their computers for running the application. Setting their preferences would make it possible to get a simple image result showing them the most suitable areas for them to stay at during their summer vacations. The idea of putting such an application online would make it more feasible for users to take advantage from it. It could be served from many organizations or companies such as tourist offices, online holiday booking systems, municipalities as well as the Hellenic Tourism Organisation itself to promote the country's tourism destinations.

One could argue: 'Why did you have to go through creating a whole script calculating the best areas when there is the simpler tool of selecting by location to get some results?' It is a very good question that makes things simpler because indeed it is possible to select the hotels which are within a certain distance from lets say beaches, and then from the result of this we can select the hotels that are within specified distance from restaurants for instance. This would definitely work but it would make our application very inflexible and not extendable. Straight distances can indeed be easier calculated. However, in the case of select by location approach we are doomed in only using straight distances. This was not an option for this research because we want to make sure that our application could actually be useful and applied in reality with the smallest possible changes. The model which was created makes use of some areas that are calculated according to a user's settings. It does not matter how these areas where created. They could be buffers, service areas from a network analysis, economic distances, time distances etc. The only thing that would have to be changed is to specify which polygons are going to be used. As a result, the way this application is developed makes it more flexible and applicable.

This application can be of help for tourists when they have to decide upon their summer vacations. But the same time the same application could also be of use for the municipality of Paros and/ or the Hellenic Tourism Organisation, as a tool used for developing tourism in certain areas which are not suitable yet. In that case, there could be a research on what is the mean preferred distance of tourists between their accommodation place and certain locations. By doing so, we could use the same model to calculate the areas that would be most profitable to have hotels. By comparing the results of the model to the existing presence of hotels, one can come to a conclusion about whether new hotels need to be built, and in which area. Of course this application would, in this case, only be part of this development possibility since there are a lot of other factors that need to be taken seriously into consideration for such a decision such as the number of tourists and the current hotel capacities and more.

REFERENCES

Arianoutsou M. et all, 2011: *Evaluating Post-Fire Forest Resilience Using GIS and Multi-Criteria Analysis: An Example from Cape Sounion National Park* in Environmental Management, *Greece*, Springer Science.

Beedasy J., Whyatt D., 1999: *Diverting the tourists: a spatial decision-support system for tourism planning on a developing island* in International Journal of Applied Earth Observation and Geoinformation, vol. 1, pp. 163-174, Elsevier Science B.V.

Beerli A., Martín J.D., 2004: *Tourists' characteristics and the perceived image of tourist destinations: a quantitative analysis - a case study of Lanzarote, Spain* in Tourism Management, vol. 25, No. 5, pp. 623-636, Elsevier.

Boniface B., Cooper C., 2009: *Worldwide Destinations Casebook*, Butterworth-Heinemann, Oxford.

Burrough P.A., McDonnell A., 2006: *Principles of Geographical Information Systems*, Oxford University Press, Oxford.

Camara A.S., Raper J., 1999: *Spatial Multimedia and Virtual Reality,* Taylor and Francis, London.

Carr N., 1999: *A study of gender differences: young tourist behaviour in a UK coastal resort* in Tourism Management, vol. 20, No. 2, pp. 223-228, Pergamon.

Carver S.J., 1991: *Integrating Multi-criteria Evaluation with Geographical Information Systems* in International Journal of Geographical Information Science, vol. 5, No. 3, pp. 321-339, Taylor and Francis, London.

Cooper C, Hall C.M., 2008: *Contemporary Tourism*, chapter 3: Contemporary tourists, tourist behaviour and flows, pp. 51-76, Butterworth-Heinemann, Oxford.

Decrop A., Snelders D., 2005: *A grounded typology of vacation decision-making* in Tourism Management, vol. 26, No. 2, pp. 121-132, Elsevier.

Dodgson J.S, et al., 2009: *Multi-criteria Analysis: a manual*, Communities and Local Government, London.

Dong P., 2008: *Generating and updating multiplicatively weighted Voronoi diagrams for point, line and polygon features in GIS* in Computers & Geosciences, vol. 34, No. 4, pp. 411-421, Elsevier.

Fernández D.S, Lutz M.A., 2010: *Urban flood hazard zoning in Tucumán Province, Argentina, using GIS and multicriteria decision analysis* in Engineering Geology, vol. 111, No. 1-4, pp. 90-98, Elsevier.

García-Grespo A. et al., 2009: *SPETA: Social Pervasive e-Tourism Advisor* in Telematics and Informatics, vol. 26, pp. 306-315, Elsevier.

Geneletti D., 2010: *Combining stakeholder analysis and spatial multicriteria evaluation to select and rank inert landfill sites* in Waste Management, vol. 30, No. 2, pp. 328-337, Elsevier.

Gibson H., Yannakis A., 2002: *Tourist roles: Needs and the Lifecourse* in Annals of Tourism Research, vol. 29, No. 2, pp.358-383, Elsevier Science Ltd., Great Britain.

Gunn C.A., Var T., 2002: *Tourism Planning*, Routledge, New York.

Hall C.M., 2005: *Tourism: Rethinking the Social Science of Mobility*, Harlow: Prentice-Hall.

Heywood I. et al, 2006: *An Introduction to Geographical Information Systems*, Third Edition, Pearson Education Limited, Essex, UK.

Hsu T.-K. et al., 2009: *The preference analysis for tourist choice of destination: A case study of Taiwan* in Tourism Management, Elsevier.

Huang Y., Bian L., 2009: *A Bayesian Network and analytic Hierarchy Process based personalised Recommendations for Tourist Attractions over the Internet* in Expert Systems with Applications, vol. 36, pp. 933-943, Elsevier.

Lin G. et al., 2002: *Examining Distance Effects on Hospitalizations using GIS: a study of three health regions in British Columbia, Canada* in Environment and Planning, vol. 34, pp. 2037-2053.

Mahasuweerachai P., Qu H, 2011.: *The Moderating Effects of Tourists' Characteristics and Novelty Seeking on the Relationships between Satisfaction, Revisit Intention and WOM*, School of Hotel and Restaurant Administration, Oklahoma State Unversity.

Malczewski J., 1999: *GIS and Multicriteria Decision Analysis*, John Wiley & Sons Inc., Canada.

Meng F., Uysal M., 2008: *Effects of gender differences on perceptions of destination attributes, motivations, and travel values: An examination of a nature-based resort destination* in Journal of Sustainable Tourism, vol. 16, No. 4, pp. 445-466.

Papanastasiou A. et. al, 2006: *Tourists' Preferences for Quality of Services: Empirical Investigation of Lesvos, Samos and Chios islands* in Tourismos: An International Multidisciplinary Journal of tourism, vol. 1, No. 2, pp. 95-101, Munich Personal RePEc Archive.

Pearce D. G, 1979: *Towards a Geography of Tourism*, Department of Geography, University of Canterbury Christchurch, New Zealand.

Pearce D. G, Grimmeau J.P., 1985: *The Spatial Structure of Tourist Accommodation and Hotel Demand in Spain* in Geoforum, volume 16, pp. 37-50, Pergamon Press Ltd., Great Britain.

Porter J.H.: *Raster & Vector GIS*, Virginia Coast Reserve LTER, retrieved from the web on 17.05.2011 (http://www.vcrlter.virginia.edu)

Rigas K., 2009: *Boat or airplane? Passengers' perceptions of transport services to islands. The example of the Greek domestic leisure market* in Journal of Transport Geography, Vol. 17, pp. 396-401, Elsevier.

Swain M.B., 1995: *Gender in tourism* in Annals of Tourism, vol. 22, No. 2, pp. 247-266, Elsevier Ltd., USA.

Thornton P.R. et. al, 1997: *Tourist group holiday decision-making and behaviour: the influence of children* in Tourism Management, vol. 18, No. 5, pp. 287-297, Elsevier Ltd., Great Britain.

Tobler W.R., 2004: *On the first law of geography: a reply. Annals of the Association of American Geographers*, vol. 94, No. 2, pp. 304-310.

Urhausen J., 2008: *Tourism in Europe: does age matter?*, Eurostat.

Wickens E., 2002: The *sacred and the profane: A Tourist Typology* in Annals of Tourism Research, vol. 29, pp. 834-851, Elsevier Ltd., Great Britain.

Witt S.F., Martin C.A., 1987: *International tourism-demand models - inclusion of marketing variables* in Tourism Management, vol. 8, No. 1, pp. 33-40, Butterworth & Co. Ltd.

Zhang H., Huang G.H., 2011: *Assessment of non-point source pollution using a spatial multicriteria analysis approach*, Environmental Systems Engineering Program, Faculty of Engineering, University of Regina, Regina, Saskatchewan, S4S 0A2 Canada

Zhu, X. et. al, 2005: *A GIS-Based Multi-Criteria Analysis Approach to Accessibility Analysis for Housing Development in Singapore*, Proceedings of SSC 2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute, Melbourne.

Κουτσόπουλος Κ. 2005: Γεωγραφικά Συστήματα Πληροφοριών και Ανάλυση Χώρου, Εκδόσεις Παπασωτηρίου, Αθήνα.

Κυριαζή Ν., 2005: *Η Κοινωνιολογική Έρευνα: Κριτική Επισκόπηση των Μεθόδων και των Τεχνικών*, Θ΄ έκδοση, Ελληνικά Γράμματα, Αθήνα.

Μπομπότη Ε., 2008: *Χαρτογράφηση του δικτύου μεταφορών για τη νήσο Πάρο με την αξιοποίηση των τεχνολογιών Συστημάτων Γεωγραφικών Πληροφοριών (GIS) και Συστήματος Παγκοσμίου Εντοπισμού Θέσης (GPS)*, Bachelor degree Dissertation, Harokopio University of Athens.

Χαλκιάς Χ., 2002: *Οργάνωση Γεωγραφικών Πληροφοριών για τα Ελληνικά Νησιά με την αξιοποίηση σύγχρονων Τεχνολογικών Εργαλείων*, Γεωγραφίες, Τεύχος 4, σ. 62-95, Εξάντας, Αθήνα.

Χαλκιάς Χ, 2006: *Όροι και Έννοιες Επιστήμης Γεωγραφικών Πληροφοριών*, Εκδόσεις ΙΟΝ, Αθήνα.

WEBSITES visited up until 30.07.2011

www.lufthansa.com

www.daarmoetikzijn.nl

www.tripadvisor.co.uk

www.easytravel.gr

www.polostours.gr

http://visit-cyclades.com

www.greecelogue.com

http://collegelifestyles.org

Paros Antiparos, Touring and Hiking Map, Topo Aegean, Cyclades, 1:40.000, Anavasi, Athens, Greece

Municipality of Paros (www.paros.gr)

Hellenic Organisation for Nature Protection (www.eepf.gr)

Blue Flag Programme (www.blueflag.org)

Blue Star Ferries (www.bluestarferries.gr)

https://greece-ferries.forth-crs.gr

www.cyclades.info

http://www.parosisland.com

Hellenic Seaways (www.hellenicseaways.gr)

Χρυσός Οδηγός (Greek Yellow Pages) (www.xo.gr)

ESRI Webhelp (http://webhelp.esri.com)

E-Education Institute (https://www.e-education.psu.edu/)

Monroe County (http://www.monroecounty.gov)

The Cornerstones of a Functioning GIS, GIS Cookbook (http://www.cookbook.hlurb.gov.ph/book/export/html/6)

URBAGRAM. (www.urbagram.net)

Wikibooks (http://en.wikibooks.org)

Yacht Charters (www.yacht-charters.gr)

INTERVIEWS AND TOURIST OFFICES

Kalamboka P., personal communication, March 10, 2011, Hotel Sea View, Dryos, Paros.

Kronos Holidays Ltd., 13 Koundouriotou st., Thessaloniki, Greece.

Polizas Travel Services, 5 L. Iassonidou st., 54635 Thessaloniki, Greece.

APPENDIX 1: TRANSLATION OF GREEK BOOKS

Κουτσόπουλος Κ. 2005: Γεωγραφικά Συστήματα Πληροφοριών και Ανάλυση Χώρου, Εκδόσεις Παπασωτηρίου, Αθήνα.

Author: Koutsopoulos K.
Title: Geographical Informaiton Systems and Spatial Analysis.
Publisher: Papasotiriou
Description: This book talks about the basic concepts of GIS and Spatial Analysis. It is divided in three major parts. The first one includes an introduction to GIS, basic concepts of spatial analysis and GIS as well as the way that reality can be presented through a GIS. The second part is about GIS processes. In more detail it refers to problem definition, input, management, analysis and output of geographical data. The third and last part has to do with spatial analysis methods such as point and linear distribution, as well as conclusions.

Κυριαζή Ν., 2005: *Η Κοινωνιολογική Έρευνα: Κριτική Επισκόπηση των Μεθόδων και των Τεχνικών*, Θ΄ έκδοση, Ελληνικά Γράμματα, Αθήνα.

Author: Kiriazi N.
Titile: The Sociological Research: Critical Overview of Methods and Techniques.
Publisher: Ellinika Grammata
Description: This is a book that generally talks about how to gather information for a research.

Μπομπότη Ε., 2008: *Χαρτογράφηση του δικτύου μεταφορών για τη νήσο Πάρο με την αξιοποίηση των τεχνολογιών Συστημάτων Γεωγραφικών Πληροφοριών (GIS) και Συστήματος Παγκοσμίου Εντοπισμού Θέσης (GPS)*, Bachelor degree Dissertation, Harokopio University of Athens.

Author: Boboti E.
Bachelor Dissertation Titile: Record and Mapping of Paros' transportation network using Geographical Information Systems and the Global Positioning System.
Dissertation's full English abstract:
Mapping geographic entities is essential in managing space whether those entities are of natural substance (such as stream networks) or man-made (for example a road network).

The last forty years a number of technological innovations have been developed which are widely used for collecting, processing as well as visualizing geographic data. The most common innovations of that kind are the Global Positioning System (GPS) and Geographic Information Systems (GIS). Especially in Greece it is quite hard to obtain updated maps of isolated areas such as small islands.

In this dissertation there will be an attempt in recording, analyzing and presenting the road network of Paros Island. We will mention the nature of the above mentioned technologies as well as the way they both helped us go through this application.

In the first chapter we are going to report the basic concepts of the GPS and the way this system works. It is of great importance to be aware of the way the GPS works because that way we will be able to know the presuppositions under which this system will provide us with the most accurate and useful information. It is necessary to know the compartments which should be used when using the GPS as well as the settings that must be adjusted. Furthermore, one ought to be aware of the proper coordinate system which ought to be applied in a particular area. The reason for this is because most of the time we have to combine a various number of data coming from different types of sources such as maps, digital imagery etc. where a coordinate system has already been used.

The second chapter deals with the Geographic Information Systems (GIS). GIS happen to be the tool for representing geographic data which have somehow been collected. For this reason we ought to know as many aspects of them as possible in order to obtain a great deal of useful information. Digital Terrain Models (DTMs) are a very good example. Importing some essential data such as contour lines together with their values will provide us with the ability of creating a Digital Elevation Model which can be used to make further surface analysis. But we will not be able to manage any kind of data unless we are aware of the sources from which we can get them like the GPS or satellite imagery. Last but not least, data quality is of great importance to any application. The more accurate data we have, the best quality will we have in our final results and calculations.

On the last part of the dissertation there is a case study of Paros Island in Greece. In this chapter we will see the benefits from using both the GPS and GIS. Here we use the GPS for collecting the most accurate data possible for locating the island's road network. After transferring the data to the GIS there has been an attempt in analyzing the road network. Using various tools we managed to come up with some useful information on which we can rely in order to take crucial decisions. Through the present dissertation one can have a good grip of how the two above mentioned technological innovations can be combined for creating maps and analyzing geographical data.

Χαλκιάς Χ, 2006: *Όροι και Έννοιες Επιστήμης Γεωγραφικών Πληροφοριών*, Εκδόσεις ΙΟΝ, Αθήνα.

Author: Chalkias C.
Title: Concepts in Geographical Information Science.
Publisher: ION
Description: This book contains a list of explanations of concepts and terminologies within the field of Geographical Information Science.

Χαλκιάς Χ., 2002: *Οργάνωση Γεωγραφικών Πληροφοριών για τα Ελληνικά Νησιά με την αξιοποίηση σύγχρονων Τεχνολογικών Εργαλείων*, Γεωγραφίες, Τεύχος 4, σ. 62-95, Εξάντας, Αθήνα.

Author: Chalkias C.
Title: A Primary Approach to Geographical Information Integration for the Greek Islands with Modern Technologies.
Journal: Geografies
Publisher: Exandas.
Summary: Nowadays, there is a need for efficient management of fragile geographical regions. Greek Islands are such regions, s they are characterised by environmental vulnerability and development problems. Advanced modern technologies provide many useful tools for geographical information management and decision making. The main aim of this study is to design and evaluate a geographical information management system for Greed island as well as to produce significant geographical information with the use of this technology. This geographical information was created using cartographical characteristics, which are mainly quantitative, are distinguished in the following main categories according to

their origin: a) various cartographic measurements produced from basic GIS layers, b) calculations extracted from Digital Elevation Models, c) measurements produced from statistical tables processing, d) products of analytical procedures, and e) proposed measurements, critical combinations – indices (e.g. general relief indicator, maximum distance, shape factor etc.). Finally, we conclude with resulting remarks about the functionality and the potential uses/ extensions of the system.

APPENDIX 2: QUESTIONNAIRE

WHAT IS MORE IMPORTANT FOR YOU WHEN YOU GO FOR SUMMER HOLIDAYS?

RESTRICTIONS:
Please fill in this questionnaire **only** if you do not travel with children, you do not travel alone and your age is between 25 and 64 years old.
INSTRUCTIONS:
Please put in order the following summer vacation activities by assigning numbers to them, with number 1 being the most important. If there is something that is not included in the list but it is still very important to you, please add it at the end of the list and include it to your numbering. In case there is an activity that you are not at all interested in, leave it out of the list.

__ Going to the beach

__ Going to a restaurant/ tavern

__ Visiting monument, traditional churches/ monasteries

__ Doing sports

__ Visiting archaeological sites

__ Going to bars/ cafeterias for a drink

__

__

__

Thank you very much!

APPENDIX 3: HOTEL MANAGER PHONE INTERVIEW

Can you estimate where most of your customers come from according to your room bookings in the period between the first of June and the 31st of September?

"Usually our customers are mostly from Greece. However there are also quite some customers from European countries, mainly from France, Germany, England and Italy. There are also some non-European visitors, but they would definitely be the minority of the customers and they mainly come from the USA.

Do your customers ask you information about whether something specific is close to your hotel? If yes what is most important for them to have close to their accommodation?

"Of course they do ask. A lot of times they ask at the time of the booking if this is done via a phone conversation but it also happens that they also ask only once they have arrived at the island. What they are mostly interested in includes beaches, taverns/ restaurants, super markets, bars and sports, mainly wind surfing and horse-back riding. Sometimes they also ask whether it is possible to reach the island by airplane. Obviously this is asked from customers that cave not arrived to Paros yet."

APPENDIX 4: TOURIST OFFICE INTERVIEW

Could you please tell us if you have some summer packages for holidays to Cycladic islands? If this is the case what is usually included in the programme?

"When it has to do with summer vacation in Greece we usually include in our offers only the hotel bookings and the transportation to the destination and back. We most of the times do not offer a specific programme to follow throughout the day. Our customers have to decide themselves what they want to do. Easter holidays are more scheduled and our packages include more activities since this period is more structured because of Church activities that are extremely popular for Greek people. Another case where we provide our customers with a detailed schedule is when they go abroad, to another country".

And what about the foreign tourists? Do they also ask you about Greek destinations?

"We do not have so many foreign customers. They usually refer to their local tourist offices which provide them with destinations and also with more detailed information of what activities they can join at their destination."

In the case that you provide some more detailed schedule about a certain summer vacation destination in Greece, what is it that tourists are mostly interested in?

"I would say that most of the times that customers are interested in visiting something specific then they usually ask for archaeological sites, monuments and thematic parcs."

APPENDIX 5: SCRIPT CODE

```
#MCA Intersection

#Function for recursive intersection
#Intersect "Layer" with all Layers in "ResultList" one by one
#Return a list with all the intersections

def RecIntersect(LayerF,ResultListF):
    #show what the current Layer is
    arcpy.AddMessage("Current Layer: " + LayerF)

    #Add current Layer to ResultList
    ResultListF.append(LayerF)

    #Copy the ResultList to a temporary list so that the working list does not expand
    TempListF = ResultListF[:]

    #Intersect all "ResultLayers" from "ResultList" with "Layer"
    #and add the result to the temporary list
    for ResultLayerF in ResultListF:
        #make the intersection only if the intersecting
        #layers are not the same
        if LayerF != ResultLayerF:
            IntersectionResultF = LayerF + ResultLayerF
            arcpy.Intersect_analysis([LayerF,ResultLayerF],IntersectionResultF)

            arcpy.AddMessage("Intersection Result: " + IntersectionResultF)

            #proceed with dissolving erasing and overwriting
            #layers only if the intersection has a result
            if arcpy.GetCount_management(IntersectionResultF).getOutput(0) != '0':

                #Dissolve the result of the intersection
                DissolveResult = "Dslv" + IntersectionResultF
                arcpy.Dissolve_management(IntersectionResultF,DissolveResult,"level")

                #Update the level number in the intersected layer by
                #setting a cursor on the tables and adding the level values
                LayerF_cursor = arcpy.SearchCursor(LayerF,"","","level")
                ResultLayerF_cursor = arcpy.SearchCursor(ResultLayerF,"","","level")
                DissolveResult_cursor = arcpy.UpdateCursor(DissolveResult,"","","level")

                #set the cursor to the next row, read the level number and store it to a variable
                row = LayerF_cursor.next()
                LayerF_level = row.getValue("level")

                row = ResultLayerF_cursor.next()
                ResultLayerF_level = row.getValue("level")

                #add up the levels of the intersected layers and
```

```
            #store the result to the level of the intersection
            sum_level = LayerF_level + ResultLayerF_level

            row = DissolveResult_cursor.next()
            row.level = sum_level
            DissolveResult_cursor.updateRow(row)

            #delete cursor to unlock files
            del row
            del DissolveResult_cursor
            del LayerF_cursor
            del ResultLayerF_cursor

            #Erase the dissolved result of the intersection from both the intersected layers.
            EraseResult1 = "in" + LayerF + "out" + DissolveResult + "erase"
            EraseResult2 = "in" + ResultLayerF + "out" + DissolveResult + "erase"

            #Replace the intersecting layers with the outcomes of the erase function.
            arcpy.Erase_analysis(LayerF, DissolveResult,EraseResult1,"0,5")
            arcpy.Erase_analysis(ResultLayerF,DissolveResult,EraseResult2,"0,5")
            arcpy.Delete_management(LayerF)
            arcpy.Rename_management(EraseResult1,LayerF)
            arcpy.Delete_management(ResultLayerF)
            arcpy.Rename_management(EraseResult2,ResultLayerF)
            #add the final intersection result to the TempList
            TempListF.append(DissolveResult)


    #return the result to the temporary list
    return TempListF

import arcpy, os

from arcpy import env

#create a list where all the initial layers will be filled in
LayerList = []

#check that the buffer layer exists and append it to the list
beaches_wholepath = arcpy.GetParameterAsText(0)
if not beaches_wholepath:
    arcpy.AddMessage('there is no beach preference specified')
else:
    beaches = os.path.basename(beaches_wholepath)
    LayerList.append(beaches)
    workspace_path = os.path.dirname(beaches_wholepath)

sites_wholepath = arcpy.GetParameterAsText(1)
if not sites_wholepath:
    arcpy.AddMessage('there is no sites preference specified')
else:
```

```
    sites = os.path.basename(sites_wholepath)
    LayerList.append(sites)
    workspace_path = os.path.dirname(sites_wholepath)

parks_wholepath = arcpy.GetParameterAsText(2)
if not parks_wholepath:
    arcpy.AddMessage('there is no parks preference specified')
else:
    parks = os.path.basename(parks_wholepath)
    LayerList.append(parks)
    workspace_path = os.path.dirname(parks_wholepath)

restaurants_wholepath = arcpy.GetParameterAsText(3)
if not restaurants_wholepath:
    arcpy.AddMessage('there is no restaurants preference specified')
else:
    restaurants = os.path.basename(restaurants_wholepath)
    LayerList.append(restaurants)
    workspace_path = os.path.dirname(restaurants_wholepath)

#set the path of the final file's location
storage_path = arcpy.GetParameterAsText(4)

#set the workspace and allow overwriting of files
env.workspace = workspace_path

arcpy.AddMessage(workspace_path)

env.overwriteOutput = True

#create the ResultList list
ResultList = []

#show what is in the LayerList list in a message
arcpy.AddMessage(LayerList)

for Layer in LayerList:
    #For each "Layer" in "LayerList" intersect the "Layer" with all Layers in
    #"ResultList" one by one and store all the intersections again in ResultList
    ResultList = RecIntersect(Layer,ResultList)

#merge all the final layers into one layer
arcpy.Merge_management(ResultList,storage_path + "/merged_final_result")

#give the result as an output to the model
arcpy.SetParameter(5,storage_path + "/merged_final_result")
```

APPENDIX 6: THE ROLE OF DISTANCE IN HUMAN DECISION MAKING

In this chapter we will elaborate how distance can influence human decisions. Since our study is mainly based on defining and calculating areas based on defined distances, we have to take a closer look and investigate the role of 'distance' as in what it exactly is (what is the definition of it), how it can be measured and of course how it can influence human decisions, and especially in the sector of tourism. For this reason we will discuss these issues in the following subchapters.

1.1 DISTANCE DEFINITION

There are many definitions for explaining what distance is. However, it is very complicated to find scientific literature about it. Even though there was an extended search for definitions in books and scientific articles, in all cases it is taken for granted that distance is a basic concept which is already known to the reader so no explanations of what distance actually take place[6]. The only sources found, which explain what distance is, are geography dictionaries, websites and scientific blogs. Therefore, we will mention some distance definitions according to the information available.

Distance is a concept that is widely used from simple everyday life to complex scientific formulas. It is one of the concepts which look like they do not need an explanation (Splung Physics Forum, retrieved: 06.10.2011). However, we need to have a definition. There are many approaches of finding what distance is. For example, in the Oxford Dictionary it is stated that distance is a 'measure of space between two points.' Another answer supports that distance is the space between two objects following the length of a certain line (Brainy Quote, retrieved: 06.10.2011). Distance is measured using different units such as meters, inches, yards, miles etc. (Oxford Dictionary). Different distance units are shown in Figure 1.1.

[6] In search for articles using the key words 'distance', 'distance definition', 'Geodesy distance', 'distance perceptions' a big number of articles came as a result (around 300.000), of which about 3.000 were related to this topic. However, in all examined articles the definition of distance was not stated, the concept was only used for further analysis in every individual study.

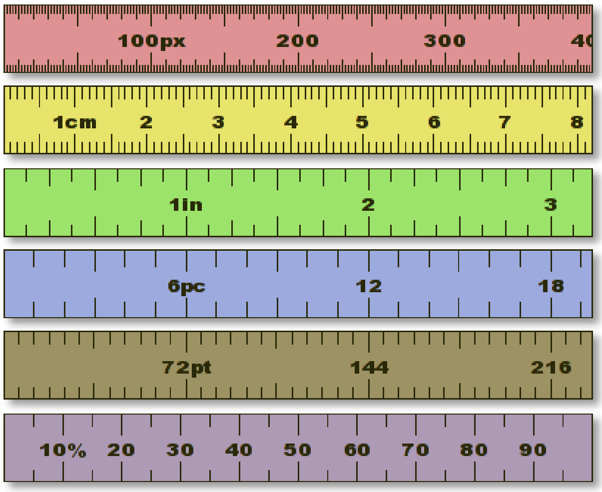

Figure 1.1: Distance in different metrics (pixels, centimetres, inches, picas, points and percentage of the ruler's length).
Source: GNOME coder.

The Oxford Dictionary of Geography states that absolute distance is expressed in physical units like meters or miles[7]. At the same time distance can also be relative, which includes any other kind of a distance measured in time, financial and other units. This is also known as cost distance. An example of relative distance is the time one needs to go from one place to another and this can vary depending on the mean of transportation. Thus, a 6-hour travel distance to a place by train could be a 90-minute travel distance by airplane. Cost distance by means of currency depends on the volume, the transport method etc. (Oxford Dictionary of Geography).

At the same time, there are more complex and scientific ways of defining what distance is. For instance, in mathematics 'distance function or metric is a generalization of the concept of physical distance. A metric is a function that behaves according to a specific set of rules, and provides a concrete way of describing what it means for elements of some space to be "close to" or "far away from" each other' (Wikipedia, retrieved: 06.10.2011).

When we talk about distance we have to make sure that we are not confusing it with displacement as it is a common mistake. According to the Splung Physics Forum, 'distance is a scalar measure of the interval between two locations

[7] The meter is the length of the path travelled by light in vacuum during a time interval of 1/299.792.458 of a second. (SI unit definition)

measured along the actual path connecting them' whereas 'displacement is a vector measure of the interval between two locations measured along the shortest path connecting them'. The difference between the two concepts is shown in Figure 1.3. The actual distance between A and B is represented by the purple line whereas their displacement is the green straight line.

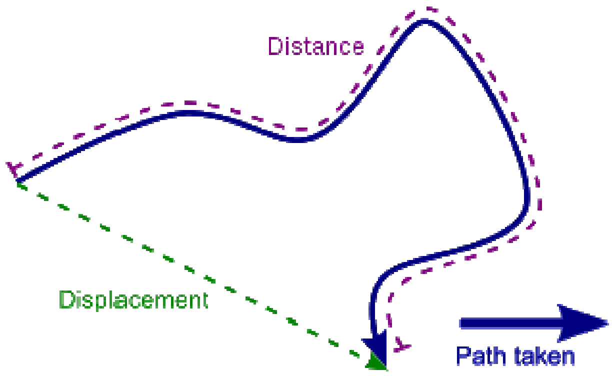

Figure 1.3: Distance (purple line) versus Displacement (green line)
Source: Wikipedia.

Distance is a concept that can be described in many different ways. However, they all have a common concept and that is that distance is the space between two spots in different places regardless how this is measured. In chapter 1.2 we will investigate the different ways that a distance can be measured.

1.2 MEASURING DISTANCES

Even though we already explained what distance is, it is important to also know how it is calculated. The simplest way of measuring distances is by using mechanic means such as measuring bars, flexible chains and measuring tape. But since measuring in this way can be complicated, electronic distance measurement devices using laser have also been widely used (Robinson A.H. et. al., 2002). In everyday discussion, distance can also be associated with a calculation based on different criteria such as using blocks as a measuring unit. One can say 'two blocks down the street' to specify how far a point is.

However, there are many other ways to calculate distances which are based on different assumptions of the measuring environment and the method used. In this subchapter we will mention some of these measurement methods including Euclidean distance, tessellation methods and projection based ones.

1.2.1 EUCLIDEAN DISTANCE

According to mathematics and geometry, Euclidean distance is the straight distance between two points and it is calculated using the Pythagorean formula (Heywood et. al., 2006) which is shown below.

$$a = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

This way of measuring works if we know the Cartesian coordinates of the two points the distance within which we want to calculate, provided that the distance measured is a straight line distance. According to the formula shown, we should have a point with x_1 and y_1 coordinates and another one with x_2 and y_2. Then we can calculate the Euclidean distance between them.

1.2.2 NETWORK BASED DISTANCES

An important way of measuring distances is through networks. There are a lot of special datasets the distance between which cannot be estimated by simply calculating the Euclidean (straight line) distance. Most of the times there is a complex connectivity between two places. Calculating distance through networks is a solution to such a problem (Greenberg J.A. et. al., 2011). For example, a distance between a restaurant and a hotel usually cannot be measured by straight line distance, it would make no sense. And that is because usually between such places are streets, other buildings and generally areas which people cannot cross. The only way is to use a road/path network to go from one place to another. For this reason in this case calculations are made based on networks.

1.2.3 FRACTIONAL DIMENSION

Another way of measuring distances is by using fractional dimension. In this method we do not measure straight lines, but lines which also form curves. Curves are sometimes so complicated to measure that it is meaningless to do so (Mandelbrot B.B, 1967). However, this method gives a solution to this problem by letting us measure along a curved line using certain length straight units as a base, like shown in Figure 1.4.

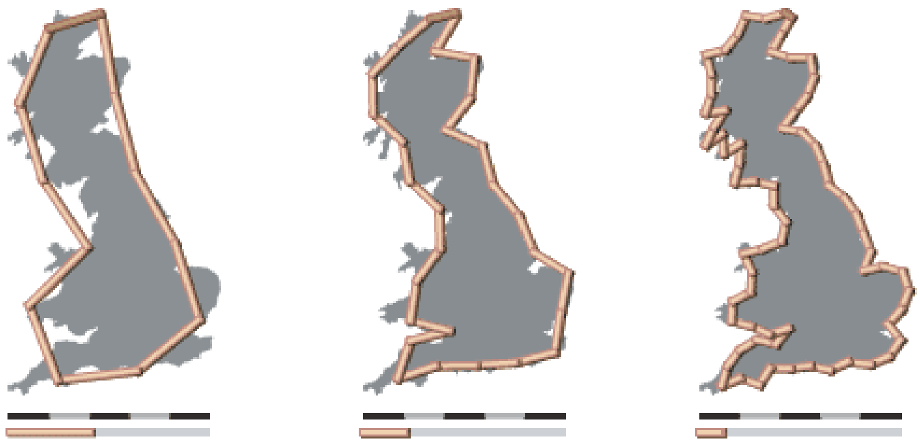

Figure1.4: Measuring distance using fractional dimension with (from left) a 200Km, a 100Km and a 50Km unit.
Source: Mathematics Illuminated (Retrieved from the web: 24.10.2011).

In this case we apply different units of fixed length along the target line and we check how many times it fits. Depending on the unit used, the results will be different. For example, the distance of Britain's coastline using diverse measuring units is shown in Table 1.1.

Unit Length (Km)	**Times Fitting[8]**	**Distance Estimation (Km)**
200	11	*2.200*
100	27	*2.700*
50	68	*3.400*

Table 1.1: Britain's coastline estimations.

As we can see, the result of the calculation strongly depends on the length of our measuring unit. The smaller the unit is, the bigger the distance estimation. That is because big units miss many curves, whereas small units fit into smaller curves, taking them into consideration. The smaller the unit, the more curves are counted in, the more precise the estimation will be.

1.2.4 PROJECTION BASED CALCULATIONS

There are times when people, and especially tourists, would like to measure distances on a map. The problem in this case is that all maps are printed based on a certain projection. As we know, by transforming geographic areas from a 3D ellipsoid shape (such as the earth) to graphs on a 2D paper (maps) using different projections, distortions can occur. These distortions (angle, direction, size and shape) alter the nature of the entities including their relative position compared to

[8] In the table we only show the integer number of times that a unit fits to the coastline. In every case there is an additional small length which is not as long as a whole unit. Therefore, the final estimation is a bit bigger than indicated.

other elements, their size relations and so on, which result in difficulties for any kind of measurement. Measuring distances can also be very challenging. There are many projection methods developed to minimize these distortions. Unfortunately, there is no projection which eliminates all types of distortions, but every one deals with a couple of them (Kraak M.J., Ormeling F., 2003).

In our case, the most suitable map projections are the equidistant ones because, according to Kraak and Ormeling, 'Equidistant map projections preserve distances between certain points'. Therefore, we can measure distances in maps with such projections with limited error. Such a projection is the Azimuthal Equidistant Projection centred to a certain point. All distances from another point to this one, in a straight line, can be correctly measured. The disadvantage is that for every combination of two places we need a projection centred to one of the two points and this is not so convenient.

Another projection which could be used to measure distances rather accurately is the UTM projection (Universal Transverse Mercator) which consists of a combination of map 'stripes' including a longitude width of 6°. They have small distortions because there is one imprint every 6°. It is a compromise but definitely not the perfect solution. In general, the smaller the area projected, the less distortions will occur.

1.2.5 TESSELATION MEASUREMENTS

In geosciences a lot of times we have to deal with tessellations which represent reality whether these are squares, triangles or octagons. Therefore, methods calculating distances on these tessellations have also been developed (Κουτσόπουλος Κ., 2005). In this subchapter we will mention a few of these methods applied in square tessellation environments.

The first way is again the Euclidean distance but based on a grid. In this case we have the source grid (source cells) which has information on the pixels from which distances will be measured (Κουτσόπουλος Κ., 2005; esri webhelp, retrieved on 25.10.2011). In grid distance calculations the distance between two cells is calculated basically by using the Pythagorean Theorem.

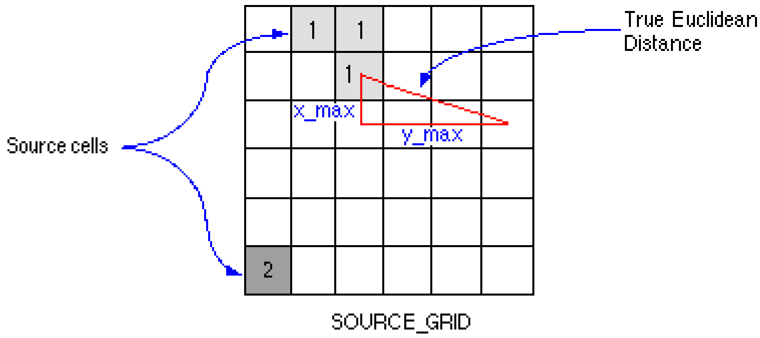

Figure: 1.5: Euclidean distance calculation on a grid.
Source: ESRI webhelp.

Figure 1.5 illustrates how these calculations are done. As we can see, the true Euclean distance between two cells is the hypotenuse of a right triangle whose other two edges are formed through the cells. By knowing the pixel size and applying the Pythagorean Theorem the true Euclidean distance can be easily calculated (ESRI webhelp). The formula for it is $a = \sqrt{b^2 + c^2}$ where a in this case is the hypotenuse and b and c are the other two edges.

Another solution in tessellations is to use the 'Manhattan' distance method. Here, distance is calculated along cell sides from one point to another (Heywood I. et al., 2006) where paths are strictly horizontal or vertical and not diagonal, as shown in Figure 1.6a.

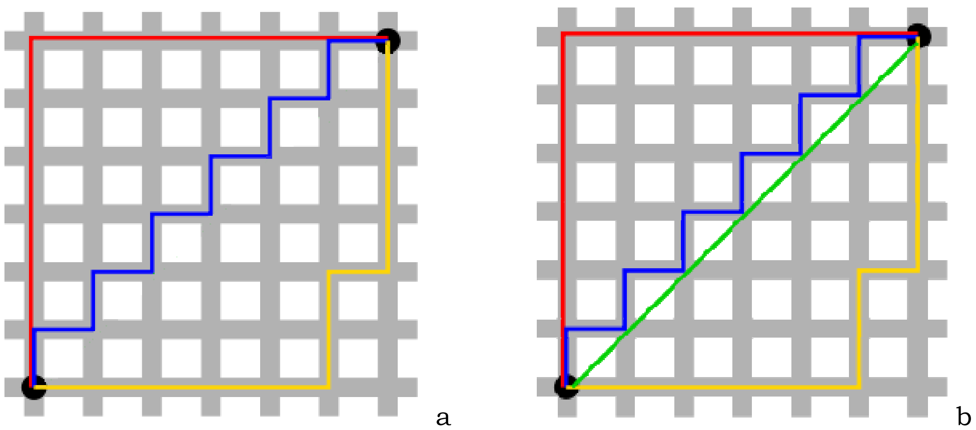

Figure 1.6: a. Manhattan distance (red, blue and yellow line) b. Manhattan and Euclidean distance (green line) compared.
Source: Wikipedia.

The read, blue and yellow lines show three equivalent Manhattan distances. Every step's length is the length of one cell. In this case all three lines give the result of 12 measuring units (pixels). This seems like an easier way, however it can cause a serious loss of information (Burrough P.A., McDonnell R.A., 2006). If we calculate the Euclidean distance for the same points (Figure 1.6b) we would find that the result would be 8,5 units which is much less than the Manhattan one. Especially when the resolution of an image is rather low, the difference in the two calculations can be enormous.

By comparing the two ways of tessellation measurement, we can conclude that the Euclidean distance is much closer to reality (when measuring straight lines) and therefore more precise. For straight line distances the Manhattan approach would diverge a lot from reality. However, it would be more useful in environments which would have such a structure that straight lines would not give precise results like in a densely populated area which is full of rectangular blocks with roads going through them.

1.2.6 SPHERE CALCULATIONS

All the above mentioned methods only apply to plane coordinate systems. Nevertheless, these solutions cannot be applied in real world where distances are measured on a curved surface like the one that earth has, having the shape of more or less an ellipsoid. However, to make calculations simpler earth is sometimes considered to be a sphere. In this case, a distance between two places along a great circle[9] is the shortest distance between these points (Robinson A.H. et. al., 2002). To calculate such a distance between two points A and B, we have to use the sphere trigonometry formula:

$$\text{Cos}(D) = (\sin(a)*\sin(b))+(\cos(a)*\cos(b)*\cos|\Delta\lambda|) \quad \text{where:}$$

D: the angle between A and B on the great circle
a,b: geographic latitudes of A and B
$|\Delta\lambda|$: the absolute value of latitude difference between A and B

Knowing the cos(D) we can also calculate the actual distance in meters between the two points using the formula:

$$\text{Distance} = \left(\frac{2\pi r}{360}\right)D$$

Where r is the average radius of the earth (6.370.997m)

[9] Great circle is the trace of the intersection of a plane with a sphere when this plane goes through the centre of the sphere.

This is a solution to measure distances on a sphere but like we already mentioned, the earth is more of an ellipsoid. If we want to have more precise results we need to apply the corresponded formula for ellipsoids. Since this is a very complex formula expanding almost a whole A4 page (see Vincenty T., 1975), it is out of the limits of this paper so we will not present it in detail. However, we can mention that the distance, based on the WGS84 ellipsoid, between Amsterdam[10] and Paros[11] is 2.321 Km, whereas the same distance calculated with the sphere formula is 2.319 Km. Even though the difference is only 2,203 kilometres, we can see how the different approaches give us different results.

1.2.7 FUNCTIONS FOR NON-NUMERICAL ATTRIBUTES

Till now, we have examined distance measurement methods which produce **numerical** attributes of distance (quantitative ratio-measurement scale). This means that the operations use standard arithmetic functions like addition multiplications and so on of numbers. Nevertheless, in many applications, there is also the need to handle **ordinal** or **nominal** attributes. Ordinal attributes might include numbers and have a hierarchical order but cannot be included in arithmetic operations because the result would not make any sense. Nominal attributes can also be text (Domingo-Ferrer J., Solanas, A., 2008). Therefore, there are some distance functions especially designed to deal with those attributes (Wilson D.R., Martinez T.R. 1997). For instance, The Value Difference Metric (VDM) and the Modified Value Difference Metric (MVDM) were developed to 'define an appropriate distance function for nominal attributes' as Wilson and Martinez state using different weighing schemes. However, these functions do not handle continuous attributes so well. For this reason, even more functions were introduced in order to also include continuous attributes, such as the Interpolated Value Difference Metric (IVDM) and the Windowed Value Difference Metric (WVDM) (Wilson D.R., Martinez T.R. 1997). These distance measurement methods are beyond the intentions and scope of this paper.

1.2.8 CONCLUSIONS

As it was discussed in this subchapter, distance can be complicated to measure because there are a lot of ways to approach the problem. One should investigate and decide upon which is the best (or a combination of them), according

[10] 52°22′23″ N, 4°53′32″ E
[11] 37°5′0″ N, 25°9′0″ E

to their needs. In our case we only took into consideration vector Euclidean calculations and network based ones because we wanted to measure straight line distances and distances through a road network. In addition, we used the Greek Grid projection which is focused on the area of Greece in order to minimize the distortions in the study area. This is a Transverse Mercator projection with the central meridian being on 24° east. The tessellation methods were not used at all because we included no grid data in our application.

1.3 DISTANCE INFLUENCE IN TOURISM DECISION MAKING

Distance can be a decisive factor in many and various parts of life activity, including tourism. A lot of times distance has a friction effect upon tourism destination selection decisions. There have been many travel decision models which include actual distance measurements between a place of origin and potential destination. However, calculating depending on actual distances is a bit risky because for every person distance is perceived in a different way. Not so much is known about the fundamentals behind the processes people follow for estimating distances (Friedman A., Montello D.R., 2006). Nevertheless, there are some attempts trying to understand these processes. Distance perception is a result of a combination of things such as an individual's senses, education, organisational structure, life experiences and imaginations. The process by which a person acquires a geographic distance, transforms it to spatial knowledge and reflects it as distance judgement is called cognition of distance. Therefore, **cognitive distance** is the distance estimated by a person based on his personally gathered information as well as his/her beliefs (Ankomah P. K., Crompton J. L., 1992; Ankomah P. K. Baker D. et. al, 1996; Crompton J.L., Kim S.-S., 2001). Cognitive distance is believed to be one of the most important criteria which contribute on tourists' destination decisions (Ankomah P. K. Baker D. et. al, 1996).

1.3.1 COGNITIVE DISTANCE PERCEPTION

Cognitive distance can be distorted some times according to each person's perspective and that can lead into miscalculations of real distance between two places and therefore a decision based on that calculation may be false. There are two schools of thoughts regarding how this information can be altered. The first school believes that the problem starts already when encoding the information in our memory whereas the second one supports that distortions occur when retrieving already stored information (Ankomah P. K., Crompton J. L., 1992).

In the first school of thought, humans are expected to memorise distances from the beginning in a wrong way. The main problem is that the distance between elements which belong to a different super-ordinate unit (large area units which include subdivisions) is not encoded correctly. For example, cities which are located in different states will cause problems in calculating the distance between each other because humans store different 'regions' in different parts of memory. This lack of proper information storage leads into overestimation and underestimation of a certain distance. To be more precise, there is a tendency of underestimating the distances between two objects which are within the same region, let's say two towns in the same county, whereas overestimations occur when the places lie within different regions (Friedman A., Montello D.R., 2006). In other words, and following the preposition suggested by Ankomah P. K. and Crompton J.L.:

'Distances between an origin and a destination that are in different super-ordinate hierarchies are likely to be overestimated, while a similar distance between an origin and a destination located within a super-ordinate hierarchy are likely to be underestimated.'

This has also been proved in the study of Friedman A. and Montello D.R. who came up with some graphic results of this theory shown in Figure 1.7.

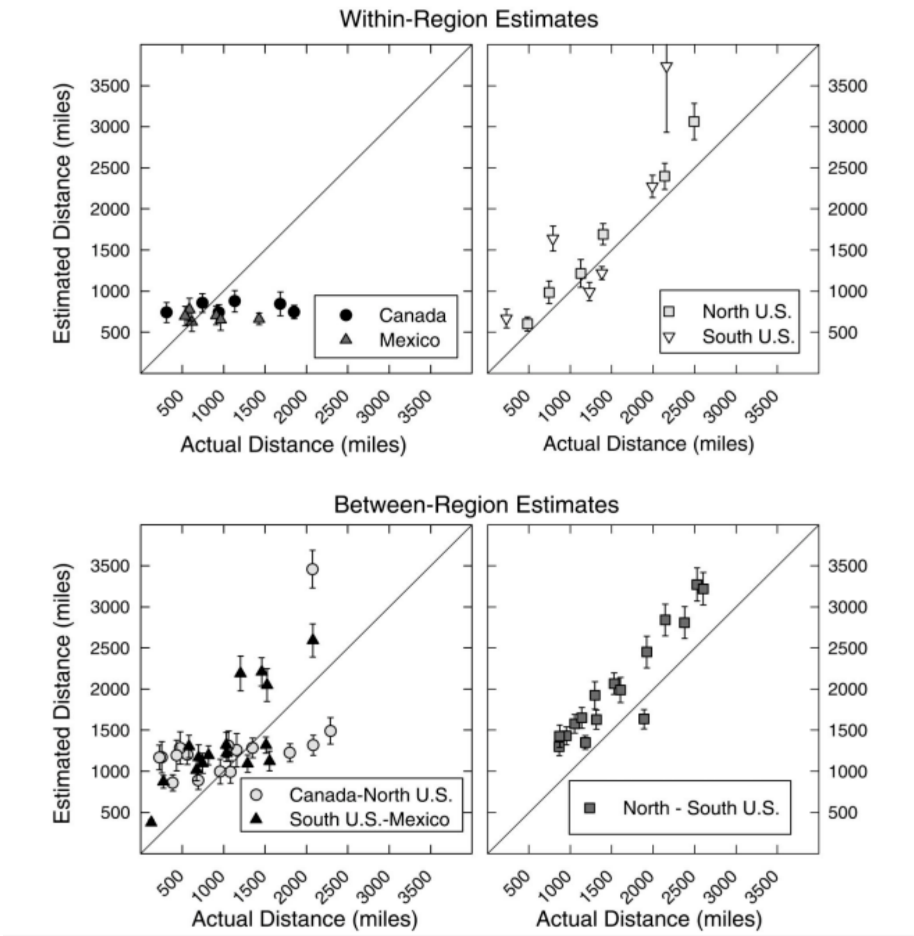

Figure 1.7: Within and between regions estimations.
Source: Friedman A., Montello D.R.

Figure 1.7 indicates that when candidates had to estimate distances between towns which where within certain regions (Canada, Mexico, North or South U.S.) they most of the times made underestimations. However, when the same candidates were asked about the distance between inter-regional cities, most of the estimations where overestimated.

All the above mentioned also apply to distances between a starting point and a tourism destination. If the destination is within the same region with the departure place then the distance will be considered to be shorter that it is in reality. In addition, if a person has a lot of information about an area then the distances between places within that area are likely to be underestimated. Furthermore, underestimations may be noticed on the distance between two points that belong to the same main route whereas the distance between places located in different main routes are likely to be overestimated (Ankomah P. K., Crompton J. L., 1992).

In the second school of belief, people estimate distances by retrieving information that has already been stored in their memory. They retrieve mind images and they scan them to come up with an estimation. The longer it takes to scan a mind image, the bigger the distance calculation will be between two points.

In Ankomah and Crompton's research there are a few studies explaining how distance estimations via a memory image process can be calculated. For example, if there are a lot of environment features retrieved from the memory then the distance estimations will again be bigger than they really are, causing overestimations. Some studies also point out that distance estimations also depend on the number of turns and intersections along the way. The more intersections and turns, the bigger the distance between two places will be mentally calculated. To make things a bit more concrete we will mention six prepositions which were stated in the same study especially for tourism.

The first one says that it is likely to have overestimations of distance calculation to destinations which are located on routes that include many landmarks whereas the opposite happens if those landmarks are fewer. The second proposition states that people who are actively taking part in travelling (such as vehicle drivers) are expected to make more accurate estimations of distances rather than those who are passively travelling (simple passengers). The third preposition supports that travellers who have seen a map of a certain area beforehand will probably underestimate a distance within the area while those who have an image from previous experience (those who have visited the place before) will rather overestimate the same distance. In the fourth preposition, a tourist will underestimate the distance towards an attractive destination, but if the destination is not so attractive, the same distance would be overestimated. The sixth preposition claims that if an area is rather flat, then the distance estimations are more accurate than in areas which are hillier (Ankomah P. K., Crompton J. L., 1992).

1.3.2 THE FACTOR OF TIME

The time one needs to travel from one place to another can also play a very important role in how a person will comprehend distance. Most of the times people prefer spending time in leisure rather than travelling (Oort C.J., 1969). Therefore, if two destinations can be reached within the same time, one might think that real distance is not such a problem anymore. By using different means of transportation (e.g. car or airplane) one can reach a destination in the same amount of time regardless if this destination is 100 or 500 Km away. Fast transportation means reduction of distance from a time point of view. That can influence the decisions a person would make when choosing a destination. It doesn't matter anymore how far in real distance a destination is as long as it takes up to a certain time to reach it. Figure 1.8 shows the railway of Bavaria in Germany presenting the distances in time.

Figure 1.8: Time distances of the railway connections in Bavaria.
Source: Deutsche Bahn.

Another influence of distance perception within this category is the socio-temporal approach. According to Matthews and Matlock, social relationships can also influence the time distance that one needs in order to get from one place to another. They made an experiment on how long it would take one to reach a certain destination in two cases: first the candidate would have to pass by people he/she already knew, and second he/she would not meet anyone familiar on their way. The results were that if a person passes by a couple of friends, it would take him longer (time-wise) to reach his destination because he/she would stop to spend some time to socialise with his acquaintance (Matthews J.L., Matlock T., 2011).

1.3.3 THE FACTOR OF COSTS

Another factor that can influence distance perception is costs, how much money a person would spend to reach a certain destination. Sometimes people think that a destination is attractive by looking at the financial costs of getting there. This may influence their decision in the end. For example, it is possible that a trip from Munich to Frankfurt and back by fast train can be more expensive than travelling from Munich to Greece and back by airplane. The difference in real distance (kilometres) is big but the cost distance (in money) is the same. So in this case, a person might consider that the 'distance' of these two destinations is the same. So the distance someone would travel in order to get to a destination in this case depends on how much money he is willing to pay for it.

1.3.4 STEPS IN DECISION MAKING

At this point it is necessary to mention the process which a tourist uses to select some possible vacation locations. Figure 1.9 briefly illustrates this process using a flow diagram.

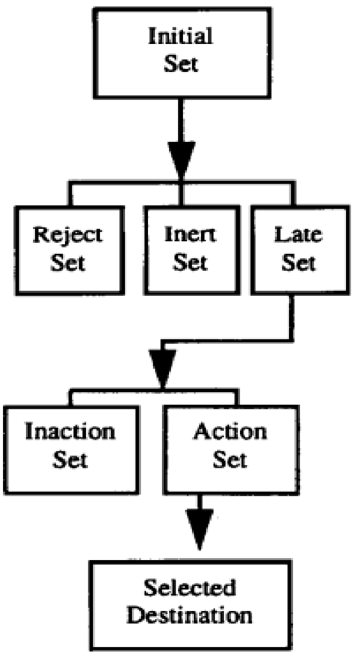

Figure 1.9: The steps for selecting a destination.
Source: Ankomah P. K. Baker D. et. al. (1996)

At the beginning, a tourist has the Initial set of destinations in mind which include all possible destinations. This set is divided in three subcategories. The Reject Set which includes the destinations that the tourist will not take into consideration at all, the Inert Set which is the sum of destinations upon which there is neither a positive nor a negative feeling and finally the Late Set where the tourist assigns the destinations which are likely to be visited. This latest set is the most important one because most of the times the tourist will decide upon a destination from this group of choices. The Late Set is again divided into two subcategories, depending on whether the tourist has gathered some information about some places (Action Set) or not (Inaction Set). Usually, the final choice will be part of the Action Set of choices (Ankomah P. K. Baker D. et. al, 1996).

Within those sets, distance estimations may vary. Figure 1.10 shows how people can miscalculate distances to locations of the same set. The most accurate estimations seem to be to places of the Late Set, and that is because these are the places that a person would be most interested in. Therefore, there are not so many overestimations. The destinations for which no real opinion exists (Inert Set) tend to be overestimated but not as much as those in the Reject Set. In the last set most of the calculations are overestimated which will probably make the person think that these destinations are a bit too far away even though they are not in reality.

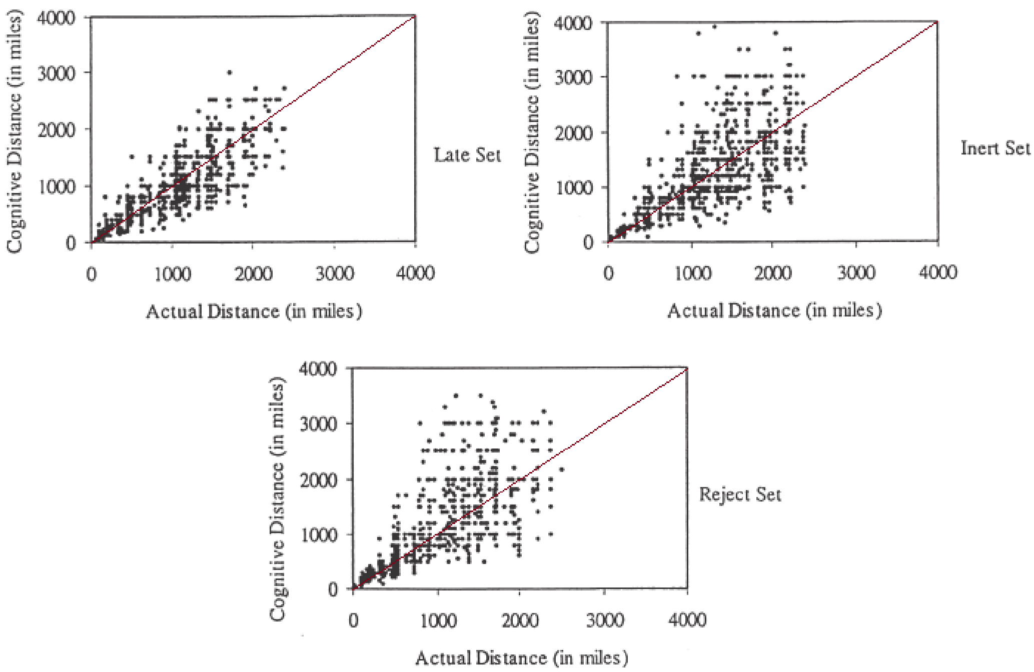

Figure 1.10: Distance calculations in decision sets.
Source: Crompton J.L., Kim S.-S. (2001)

1.3.5 CONCLUSIONS

In this part of the chapter we have shown how distance perception can be different for every person according to certain conditions under which he/she has to make an estimation. In addition, different decision phases can also influence this estimation for or against one's favour. Therefore, when building a tourist decision support system, we have to always be aware of the kinds of estimation miscalculations, which may be in real distance, time, money, which may influence the final decisions of tourists.

1.4 GENERAL CONCLUSIONS

Coming to our conclusions, we have to mention that distance can definitely influence human decision making, even if that has to do with vacation destination selections. Cognitive distance can be a major factor which may alter the classification of destinations because of over or underestimating distances. (Ankomah P. K. Baker D. et. al, 1996). The different kinds of distortions mentioned, in combination with the incorrect calculations and interpretations within different stages of decision making, are likely to lead towards different final destination decisions than one would expect. It seems that distance calculations can have diverse results based on how each person apprehends the factor of distance and his/her perception of space.

Nevertheless, through the development of various GIS tools, distances can be more accurately measured with a computer. There are many methods implemented in GI systems, like the ones explained in chapter 1.2, which can calculate lengths precisely (esri webhelp) in some milliseconds. If a tourist would have the opportunity to use such a tool for distance estimations, it could influence his decision because it might give a different result than the one that he has mentally calculated. People also tend to 'trust' the computers' results since they know that they are more accurate and closer to reality than their personal rough mind calculation. Therefore, if a person has overestimated a certain distance but in the end a geo-tool comes up with a shorter distance result, then he might reconsider including the certain place to the potential destinations. Without a geo-tool it could be a reason to exclude a potential destination. At this point one can question how cognitive distance can be influenced by geo-tools. Our practical application takes into account personal distance preferences. What it does is to help the user estimate which areas are within certain distances from certain places in a graphic representation (maps). The results are most likely different to the user's personal calculations and therefore it can change his understanding of distance as well as his decisions and therefore alter the cognitive distance perception. The way modern GI tools, like this one, can influence cognitive distance is a very interesting topic that comes as an outcome from this chapter and would need further research.

In our applied case, we did not make use of cognitive distance as input. Since we developed a special application for tourist consulting, distance calculations where already made by geo-tools and especially by using the above mentioned straight line Euclidean distance and the network based methods. Based on that, we could come up with rather precise suggestions for tourism destinations; however every individual's distance perspective was not taken into consideration.

There are many more aspects that can alter distance calculations. Like it was mentioned in chapter 1.3, time and costs can also influence the importance and view of distance and therefore also lead into different decisions regarding destination choice. There are more ways of distorting distance perception, such as temperature, energy, terrain alternations and so on, which could influence calculations and human decisions as well, however that goes beyond the limits of this chapter.

REFERENCES

Ankomah P. K., Crompton J. L., 1992: *Tourism Cognitive Distance: A Set of Research Proposition* in Annals of Tourism Research, Vol. 19, pp. 323-342, Pergamon Press Ltd.

Ankomah P. K., Baker D. et. al., 1996: *Influence of Cognitive Distance in Vacation Choice* in Annals of Tourism, Vol. 23, No. 1, pp. 138-150, Elsevier Science Ltd., Great Britain.

Burrough P.A., McDonnell R.A., 2006: *Principles of Geographical Information Systems*, Oxford University Press.

Crompton J.L., Kim S.-S., 2001: *Influence of Cognitive Distance in Vacation Choice: Research Notes and Reports* in Annals of Tourism Research, Vol. 28, No. 2, pp. 512-515, Elsevier Science Ltd., Great Britain.

Domingo-Ferrer J., Solanas, A., 2008: *A Measure of Variance for Hierarchical Nominal Attributes* in Information Sciences, Vol. 178, pp. 4644-4655, Elsevier.

Friedman A., Montello D.R., 2006: *Global-Scale Location and Distance Estimates: Common Representations and Strategies in Absolute and Relative Judgements* in Journal of Experimental Psychology: Learning, Memory and Cognition, Vol. 32, No. 3, pp. 333-346, The American Psychological Assosiation.

Greenberg J.A. et al., 2011: *Least Cost Distance Analysis for Spatial Interpolation* in Computers and Geosciences, Vol. 37, pp. 272-276, Elsevier.

Heywood et al., 2006: *An Introduction to Geographical Information Systems*, Third Edition, Pearson Education Ltd, England.

Kraak M.J., Ormeling F., 2003: *Cartography: Visualization of Geospatial Data*, Second Edition, Pearson Education Ltd.

Mandelbrot B.B, 1967: *How long is the Coast of Britain? Statistical Self-similarity and Fractional Dimension* in Science, Vol. 156, No. 3775, pp. 636-638, American Association for the Advancement of Science.

Matthews J.L., Matlock T., 2011: *Understanding the Link between Spatial Distance and Social Distance* in Social Psychology, Vol. 42, No. 3, pp. 185-192, Hogrefe Publishing.

Oort C.J., 1969: *The Evaluation of Travelling Time* in Journal of Transport Economics and Policy, Vol. 111, No. 3, pp. 279-286, JSTOR.

Robinson A.H. et. al., 2002: *Στοιχεία Χαρτογραφίας*, Πανεπιστημιακές Εκδόσεις Ε.Μ.Π., Αθήνα.

Vincenty T., 1975: *Direct and Inverse Solutions of Geodesics on the Ellipsoid with Application of nested Equations*, Vol. 23, No. 176, pp. 88-93, Ministry of Overseas Development, Surrey.

Wilson D.R., Martinez T.R. 1997: *Improved Heterogeneous Distance Functions* in Journal of Artificial Intelligence Research, Vol. 6, pp. 1-34, AI Access Foundation and Morgan Kaufmann Publishers.

Κουτσόπουλος Κ., 2005: *Γεωγραφικά Συστήματα Πληροφοριών και Ανάλυση Χώρου*, Εκδόσεις Παπασωτηρίου, Αθήνα.

Χαλκιάς Χ., 2006: *Όροι και Έννοιες Επιστήμης Γεωγραφικών Πληροφοριών*, Εκδόσεις ΙΟΝ, Αθήνα.

Oxford Dictionary.

WEB REFERENCES

Brainy Quote (www.brainyquote.com)
Deutsche Bahn (www.bahn.de)
ESRI Webhelp (http://webhelp.esri.com)
GNOME coder (http://gnomecoder.wordpress.com)
Mathematics Illuminated (www.learner.org)
Oxford Dictionary of Geography
(http://www.answers.com/topic/distance#Oxford_Dictionary_of_Geography_d
Splung Physics Forum (www.splung.com)
Wikipedia (www.wikipedia.org)