University of Turku
Department of Geography and Geology

Olli Jakobsson

GIS-BASED APPROACH FOR OPTIMIZATION OF ONSHORE WIND PARK INFRASTRUCTURE ALIGNMENT IN FINLAND

MSc Thesis

Keywords: GIS, path optimization, least-cost path analysis, multi-criteria evaluation, wind energy, infrastructure alignment

Turku 2016

UNIVERSITY OF TURKU
Faculty of Mathematics and Natural Sciences
Department of Geography and Geology

JAKOBSSON, OLLI: GIS-based approach for optimization of onshore wind park infrastructure alignment in Finland

MSc Thesis, 112 p.
40 ECTS
Geography
February 2016

Wind power is a rapidly developing, low-emission form of energy production. In Finland, the official objective is to increase wind power capacity from the current 1 005 MW up to 3 500–4 000 MW by 2025. By the end of April 2015, the total capacity of all wind power project being planned in Finland had surpassed 11 000 MW. As the amount of projects in Finland is record high, an increasing amount of infrastructure is also being planned and constructed. Traditionally, these planning operations are conducted using manual and labor-intensive work methods that are prone to subjectivity.

This study introduces a GIS-based methodology for determining optimal paths to support the planning of onshore wind park infrastructure alignment in Nordanå-Lövböle wind park located on the island of Kemiönsaari in Southwest Finland. The presented methodology utilizes a least-cost path (LCP) algorithm for searching of optimal paths within a high resolution real-world terrain dataset derived from airborne lidar scannings. In addition, planning data is used to provide a realistic planning framework for the analysis. In order to produce realistic results, the physiographic and planning datasets are standardized and weighted according to qualitative suitability assessments by utilizing methods and practices offered by multi-criteria evaluation (MCE). The results are presented as scenarios to correspond various different planning objectives. Finally, the methodology is documented by using tools of Business Process Management (BPM).

The results show that the presented methodology can be effectively used to search and identify extensive, 20 to 35 kilometers long networks of paths that correspond to certain optimization objectives in the study area. The utilization of high-resolution terrain data produces a more objective and more detailed path alignment plan. This study demonstrates that the presented methodology can be practically applied to support a wind power infrastructure alignment planning process. The six-phase structure of the methodology allows straightforward incorporation of different optimization objectives. The methodology responds well to combining quantitative and qualitative data. Additionally, the careful documentation presents an example of how the methodology can be evaluated and developed as a business process. This thesis also shows that more emphasis on the research of algorithm-based, more objective methods for the planning of infrastructure alignment is desirable, as technological development has only recently started to realize the potential of these computational methods.

Keywords: GIS, path optimization, least-cost path analysis, multi-criteria evaluation, wind energy, infrastructure alignment

TURUN YLIOPISTO
Matemaattis-luonnontieteellinen tiedekunta
Maantieteen ja geologian laitos

JAKOBSSON, OLLI: Tuulivoimapuiston infrastruktuurilinjausten optimointi paikkatietomenetelmien avulla Suomessa

Pro gradu -tutkielma, 112 s.
40 op
Maantiede
Helmikuu 2016

Tuulivoima on nopeasti kehittyvä, vähäpäästöinen energiatuotantomuoto. Suomen virallinen tavoite on kasvattaa tuulivoimakapasiteettiaan nykyisestä 1 005 MW:sta 3500–4000 MW:n asti vuoteen 2025 mennessä. Huhtikuun 2015 lopussa Suomeen suunniteltavien tuulivoimahankkeiden kokonaiskapasiteetti oli ylittänyt 11 000 MW:n rajan. Samalla kun tuulivoimahankkeiden määrä Suomessa on ennätyskorkea, suunnitellun ja rakennetun infrastruktuurin määrä on kasvussa. Yleisesti infrastruktuurin suunnittelussa hyödynnetään manuaalisia ja työläitä menetelmiä jotka ovat alttiita subjektiivisuudelle.

Tämän tutkielman tavoitteena on kehittää paikkatietomenetelmiin pohjautuva metodologia optimaalisten reittien löytämiseksi osana maalle rakennettavan tuulivoimapuiston infrastruktuurilinjausten suunnittelua. Tutkimusalue on Nordanå-Lövbölen tuulivoimapuisto Kemiönsaarella Lounais-Suomessa. Tutkimus perustuu korkearesoluutioisen laserkeilauksella tuotetun maastoaineiston ja kustannustehokkaan reittianalyysialgoritmin (LCP, least-cost path) käyttöön optimaalisten reittien paikantamiseksi. Lisäksi tutkimuksessa käytetään tuulivoimapuiston kaavoitukseen liittyviä aineistoja. Realististen tulosten tuottamiseksi maasto- ja kaavoitusaineisto yhdenmukaistetaan ja painotetaan laadullisen sopivuusarvioinnin perusteella, hyödyntämällä monikriteeriarviointimenetelmiä (MCE, multi-criteria evaluation). Tulokset esitetään skenaarioina, jotta niiden käyttökelpoisuutta voidaan arvioida suhteessa suunnittelun tavoitteisiin. Tutkimuksen metodinen kulku dokumentoidaan prosessimallinnuksen (BPM, business process management) avulla.

Tulokset osoittavat, että paikkatietomenetelmäpohjainen, LCP- ja MCE-menetelmiä hyödyntävä metodologia kykenee tehokkaasti etsimään ja osoittamaan laajoja, 20–35 kilometrin mittaisia, optimointitavoitteisiin perustuvia reittiverkostoja tutkimusalueella. Maastoaineiston korkea resoluutio mahdollistaa objektiivisemman ja yksityiskohtaisemman reittisuunnitelman luomisen. Tutkielma osoittaa, että esitettyä metodologiaa voidaan käytännön tasolla hyödyntää tuulivoimaloiden infrastruktuurilinjausten suunnittelussa. Metodologian kuusivaiheinen rakenne mahdollistaa sen soveltamisen erilaisia optimointitavoitteita sisältävissä tehtävissä. Metodologia kykenee sujuvasti yhdistämään määrällisen ja laadullisen aineiston. Lisäksi, kattava dokumentointi esittää, kuinka metodologiaa voidaan tarkastella ja kehittää liiketoimintaprosessin näkökulmasta. Tutkimus osoittaa, että olisi suositeltavaa jatkaa algoritmipohjaisten ja objektiivisempien menetelmien kehittämistä infrastruktuurilinjausten suunnittelun tueksi, sillä teknologian kehitys on vasta hiljattain alkanut realisoida näiden laskennallisten menetelmien laajempaa potentiaalia.

Avainsanat: paikkatietomenetelmät, paikkatieto, reittioptimointi, monikriteerianalyysi, tuulivoima, infrastruktuurilinjaukset

Table of contents

1 INTRODUCTION

Wind power is a rapidly developing form of energy production. It is a quickly growing industry worldwide, mainly due to the large leaps made in the development of wind turbine technology during the past decades and increasing investments on renewable energy sources overall (Kaldellis & Zafirakis 2011; Lynn 2012). Wind energy is currently one of the most effective and mature ways to produce energy from renewable energy resources (Lynn 2012). The main strength of producing energy with wind power is that wind power plants produce a relatively small amount of emissions during their life cycle. According to IPCC (2014), onshore wind power has the lowest median life cycle emissions among all prominent energy supply technologies. When installed and in operation, a wind turbine does not produce any emissions, such as carbon dioxide.

There is a strong agreement among climate scientists on the anthropogenic climate change (Anderegg et. al. 2009). This has increased the pressure on governments and global organizations to adopt and endorse new policies to act against climate change. The evidence on climate change can be seen as a major catalyst for policies that encourage governments to utilize renewable energy sources at the expense of fossil fuels. According to IEA Wind (2015), at the end of 2014, wind energy provided 4,1 percent of the electricity demand among IEA Wind's member countries. IEA has predicted in 2013 that the share of wind power of global electricity production will rise up to 15–18 percent in 2050 (IEA 2013).

At the end of 2015, the total wind power capacity in Finland was at 1 005 MW, producing a total of 2 329 GWh. This accounts for 2,8 percent of the electricity demand in the country (Finnish Energy 2016). Finnish government has issued target figures specifically for wind energy production. The objective is to increase wind energy production up to 9 terawatt hours of electricity by the year of 2025. This requires a capacity increase up to 3500–4000 MW. These figures would increase the share of wind energy in electricity production to about 6–7 percent. (Tuulivoimaopas 2014). By the end of April 2015, the total capacity of all wind power projects being planned in Finland had surpassed 11 000 MW (STY 2016).

Like any other form of energy production, wind power requires infrastructure. In general, the combined costs of wind park infrastructure are in the range of 10–15 percent of the total project costs (Vaasa Energy Institute 2010). On a relative scale this figure might seem like a small proportion, but on an absolute scale the costs can reach tens of millions of euros, depending on project size. Thus, the argument for careful planning and optimization of infrastructure alignment is valid, as the potential savings might reach hundreds of thousand, even millions of euros. Additionally, as the construction of wind power infrastructure networks often cover an extensive amount of land area, the optimization process can also minimize the negative effects on the local landscape and environment.

The process of infrastructure alignment planning has traditionally been conducted by civil engineers of infrastructure companies. Most of the planning work is carried out with manual and labor-intensive work methods utilizing subjective evaluation based on thematic maps, field visits and observations. While some utilize various different planning software, a large portion of civil engineers are still using different thematic paper maps, such as contour and terrain maps describing topography and land cover classification (Liu & Sessions 1993; Rogers 2005; Picard et al. 2006). This type of manual planning can be labor-intensive and time consuming (Anderson & Nelson 2004). However, there are some more quantitative and objective methods available to minimize the manual and rather subjective practices used in the planning of infrastructure alignment. Geographic Information Systems (GIS) combined with environmental modeling offer a computational approach for the optimization of infrastructure networks. As GIS-based environmental modeling has developed rapidly during the last decades, the fast development of technology has led to a situation where modern powerful computers allow applications of computational models, such as the least-cost path (LCP) analysis for evaluating real world problems (Lee & Stucky 1998). LCP analysis has been widely utilized in the planning of line-shaped infrastructure, such as walking paths (Lee & Stucky 1998), roads (Collischonn & Pilar 2000; Yu et al. 2003; Atkinson et al. 2005; Choi et al. 2009) and power lines (Bagli et al. 2011). In some studies, LCP has been combined with MCE in order to produce realistic results (Atkinson et al. 2005; Choi et al. 2009; Bagli et al. 2011). However, a methodology that combines LCP and MCE has not yet been applied to the planning of wind park infrastructure alignment.

The study has been conducted in Kemiönsaari, on the largest island in Southwest Finland in co-operation with EFE Ab, which is the owner company of the project used in this study, Nordanå-Lövböle. It is a large project on a Finnish scale, consisting of 29 turbines within an area of 1600 hectares. A project this large requires an extensive network of infrastructure and thus it is a relevant study area for optimization of infrastructure alignment. The study area is divided into multiple clusters based on a logistic assessment. These clusters have varying structures regarding existing infrastructure and environmental features. The physiographic versatility allows a broad perspective on what type of results can be produced with the presented methodology.

In this thesis, a GIS-based methodology, utilizing the combined uses of LCP and MCE methods, is applied in the planning process of onshore wind park infrastructure alignment. While GIS is utilized as the platform for the methodology, least-cost path (LCP) algorithm is used to search for optimal paths within a real-world terrain dataset. Terrain data is derived from airborne lidar scannings, providing high resolution data for environmental modeling. Additionally, planning data is used to implement a realistic planning framework for the analysis.

In practice, this means that the planning process acts as a guideline for the analysis to create a wind power framework for the study.

The research objectives of the study are:

1. Identify optimal paths for the alignment of infrastructure networks based on multiple different optimization objectives within the study area

2. Establish a methodology for determining optimal paths to support the process of wind park infrastructure alignment planning in GIS environment

3. Develop combined uses of high-resolution terrain data in a least-cost path (LCP) analysis by utilizing multi-criteria evaluation (MCE) methods

In order to produce LCPs that correspond to planning objectives, the physiographic and planning datasets are standardized and weighted according to scenario-based suitability rules by utilizing methods offered by multi-criteria evaluation (MCE). Finally, the presented methodology is documented from a metaplanning perspective by using tools of Business Process Management (BPM).

2 THEORETICAL BACKGROUND

2.1 GIS and environmental modeling

Since the emergence of Geographic Information Systems (GIS) in 1960's, a vast array of methods have been developed to capture, store, manage, analyze and view spatial data. Even though GIS has seen huge advances in its tools and practices, it has received extensive criticism from the scientific community for its often overlooked scientific basis. This chapter briefly introduces some aspects of this scientific and theoretical discourse to give a deeper perspective towards the framework presented in this study, in which GIS and environmental modeling form a quantitative and computational platform for realistic path optimization modeling.

Throughout its evolution, GIS has received criticism for its epistemological foundation. GIS is a powerful array of quantitative geographical tools, which can be, according to some, considered to resurrect the pure spatial research within the science of geography (Sheppard 2000). As a result, GIS has become the focal point of quantitative and empiric discussions. GIS can be undoubtedly stated as quantitative and empiric, because it is strongly based on computational methods. Additionally, GIS can be utilized as a tool for managing and editing spatial databases, which are based on empiricism. Some critics have also targeted the limitations of the representations of GIS. Some have questioned the way how GIS assumes that the world can be accurately modeled with algorithms combined with translation rules, which seamlessly connect the abstract variables with the objects and phenomena of real world together (Sheppard 1995).

One prominent geographer to criticize GIS epistemology, is Taylor (1990), who has claimed that while GIS is a sufficient tool for information management, it is incapable of produce anything regarding real knowledge and while it is only concerned with factual information, the possibilities for profound analyses are minimal. Goodchild (1991) has responded to this criticism by suggesting that GIS is designed to be used not as a substitute for knowledge, but rather alongside it. Goodchild admitted that GIS without geography is a dangerous and naïve empiricism and GIS is most useful when it is "guided by people trained in the nature of geographical phenomena" (Goodchild 1991:336). He even distanced GIS from quantitative geography, referring to cartography, with its generalization and fuzziness. Goodchild (1991:336) also stated that GIS, unlike the pocket calculator, should not be seen only as a tool, but rather as a way to "provoke profound geographical thoughts".

According to Maguire et al. (1991), the rapid development achieved in GIS has possibly had a negative effect on the analysis and definition of GIS. Maguire et al. suggest that the commercial nature of GIS and the ways in which GIS has been developing within different sciences and disciplines has led to a situation where these issues have been blurred because they have never been thoroughly debated or discussed in more detail. It is evident, that GIS as a tool and a process is developing rapidly due to the great advances in computer technology and investment environment.

Similarly to GIS, environmental modeling is a relatively new research area. Both of these research areas have made huge progress during the last couple of decades and are expected to have a large impact on how Earth's resources are managed and utilized (Waters 2002). According to Goodchild (1992), the main arguments for the use of GIS in the 1990s were still the same as they were in the 1960s: in environmental modeling and policy development.

Nowadays geographers have access to a vast array of tools and methods, which can be used to manage spatial data. According to Haggett & Chorley (1967), this is a result of a long-term historical development, during which geographers started developing a tendency to focus more on studying the spatial characteristics and topology- and geometry-related abstract characteristics of phenomena, instead of studying the actual phenomenon itself. Kemp (1992) has argued, that this topological-geometrical framework describing reality is directly related to the data models that GIS is offering. She claims that the growing amount of individually created interfaces which link reality and spatial data act as a proof and validation for this sort of development. Kemp feels that these kind of interfaces which rely on topological-geometrical frameworks are being developed and created by multiple professionals in many different projects.

Kemp (1992) sees significant advantages in merging GIS into a spatial data model that represents the environment. One of her strongest arguments is that the combination of GIS and environmental modeling enables to introduce the models as representations of reality. Otherwise, the models would just be plain, independent elements in a GIS database. Brimicombe (2010) has introduced an interesting concept *Solution space*, which contains three parts: 1) environmental-related engineering work to produce designs of utilitarian and realistic solutions, 2) simulation of models as a computational method and 3) GIS as a new technology, which can be used to understand the spatial dimensions of environmental issues. Brimicombe believes that the understanding and even managing of natural processes is possible through successful quantitative simulation. According to him, this concept of *Solution space* may lead to more advanced environmental modeling and additionally ecologically more sustainable solutions.

Already at the time of emergence of GIS in 1960s, it was understood, that there is a new paradigm emerging in the field of environmental modeling. The possibilities offered by spatial data and the functionalities of GIS were to lead science towards a whole new understanding of environment, new models and formulas and "new patterns of searching the real world" (Haggett & Chorley 1967). According to Harvey (1969) the new paradigm was not about some revolution of geographical information systems. Instead, according to him, it was about studying the logic behind geographical information systems and formalizing the methods, which could be utilized by geographers to spatially organize available data masses. Harvey's logic could be interpreted in a way, in which one should primarily try to understand the relationship between the shapes described in the data and the real characteristics of phenomena instead of focusing on arguing about how relevant the described data actually is.

Harvey (1969) states that unambiguous definition of a model is extremely challenging due to the various functionalities of models. He nonetheless phrases that a model can be considered as a formalized expression of theory. According to Chorley (1964), a model can be used to create a theory of the real world only when part of the reality has been successfully mapped into this model. Additionally, according to Chorley, the researcher has to succeed in avoiding too dramatic generalization of information in the abstraction phase and to be able to carry out a precise interpretation of the model into real world terms. Perhaps the most challenging problem of models is a linguistic one: their operation is based on artificial, abstract language systems, which have no empirical content or a relevant meaning (Harvey 1969). Harvey argues that the power of computer languages lies in the ability to link these abstract linguistic elements to real world phenomena. Fundamentally it is a question of how to build a bridge between a model and a representation, for example between mathematical formulas or programming languages and intelligible human language. A map, for example, is an important tool to understand spatial structures. According to Harvey (1969), geographers find maps as a special language created exclusively for spatial information.

Kemp (1992) approaches the challenge of linking abstraction and reality by offering a solution where abstract languages and theories can be connected with the real world by defining the abstract languages empirically. As an example she uses the gravity model of interaction between cities, where the model is linked to reality by demonstrating that cities form the bodies of the formula and their masses are symbolized by their populations.

Traditional theory of cartography may clarify the conceptual understanding of the relationship between models and their representations. Martin (1991) has formulated, similarly to Chorley (1964): if cartographical model represents the relationship between a map and reality, then a map may act as a representation of reality. In this model the cartographer's objective is to create as good as possible an approximation to an ideal map transformation while avoiding excessive information loss. According to Martin, the process of traditional, analog map creation can be modeled as a group of transformations between the real world, raw data, the map and the map image. In GIS context, Martin introduces an additional stage to his framework: a data model in which the data manipulation is conducted (Figure 1). The first transformation (T1) consists of real world data selection, for example surveying measurements. This raw data is then read in to the GIS in the second transformation stage (T2). In the third transformation stage (T3) within the GIS, multiple different data manipulation or management tools and operations are available for additional data transformation and data storage. This transformed data is then ready for presentation as output data (T4).

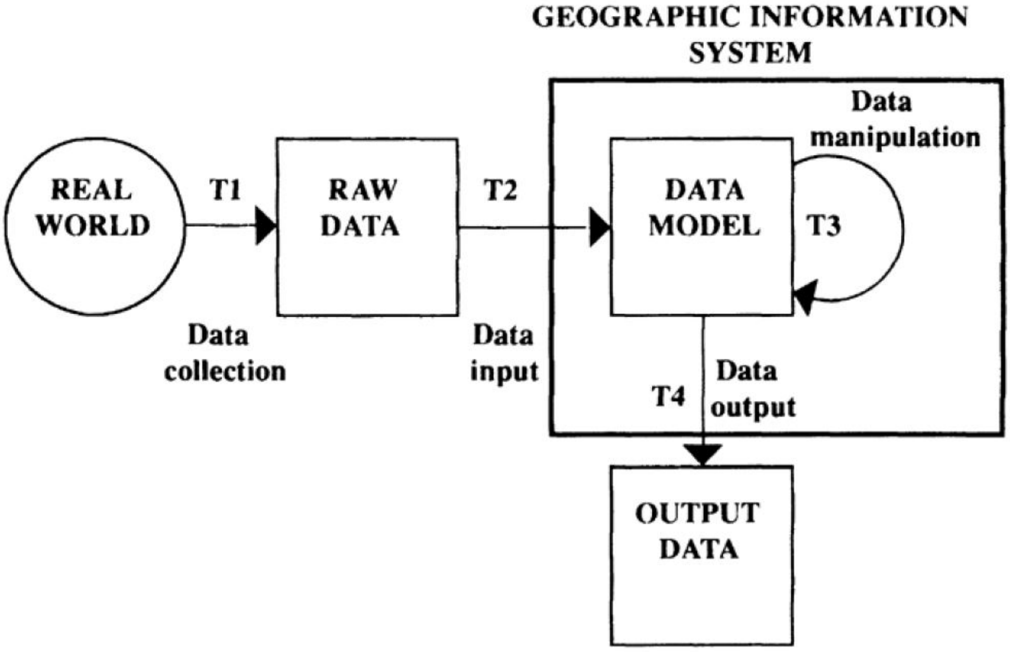


Figure 1. A transformation-based view of GIS operation (Martin 1991).

Martin argues that the significance of these transformations is that they are the single most important factor on how much information will be transmitted between different stages of the process. Martin claims that this process of transformations and representations is mainly unproblematic, and if it proves to be erroneous or flawed, it is merely due to lack of technical skills or unintentional error (Pickles 1995).

According to Malczewski (1999), the purpose of a GIS database is to represent an abstract model of reality. GIS-based applications can be only as powerful as precisely the abstract representation can model the characteristics of real-world geographical systems. Malczewski states that as a result of a conceptual debate over the aspects of real-world representations in GIS, two concepts shall be identified. The first one, *discrete object* concept, can be defined as a homogeneous area that has clear boundaries. The second, *continuous field* concept, interprets the world as a continuous field of points, and each of them contains a function that defines the location and its attributes.

Burrough (1992) has described the levels of abstraction used in the process of conceptualizing the reality and its representation in GIS databases. First level consists of conceptual models of real-world phenomena, according to which, the world can either be seen as full of defined entities with clear boundaries, or as a continuous field of spatial variation. Second level is comprised of data models of real-world phenomena. Burrough argues that just like in quantum physics, in the natural world research issues there is an existing duality between seeing the world as consisting of either fully definable objects or entities or alternatively composed of units which are incomplete, inexact and undefinable. The former, exact and defined data model is directly linked to discrete object concept and exemplified by e.g. areas or lines. The latter, inexact, incomplete and undefinable data model called *complex continua* is in turn linked to the continuous field concept and illustrated by e.g. elevation models, vegetation types or soil types.

Third level is formed by database models and their implementation. Database models are the encapsulation of the data models in which the geographical data can be saved and managed. In the case of exact data models, a database may consist of objects, attributes, relations and rules to define their behavior. This type of model is implemented by using discrete objects, such as points, lines or polygons. The field model or complex continua model, on the other hand, can be represented with discretized surfaces, continuous mathematical functions and their parameters or as continuous mathematical functions. Fourth and the last level is graphic models and their implementation. Both of the data models, the discrete object model and the field model (complex continua) are possible to be graphically implemented using either the raster or the vector data structures.

Malczewski (1999) has also described a more pragmatic approach to the process of raw data transformation through modeling into information that represents reality. According to this approach, GIS should be considered as a process, in which the main objective is to support decision making. It is essential to recognize the objectives of the task during the whole cycle of the process. Malczewski has studied the representations of GIS-based information and he has investigated the relationship between geographical data, geographical information and decision making. He sees geographical data as raw material. To utilize this information for example in decision making, it has to be processed and transformed into relevant geographical information. The main argument is, that data transformation makes the data more valuable, and thus it becomes more useful information to its users, for example decision makers. Following the same logic, we may imagine a situation in which a researcher inputs geographical raw data into a mathematical model. The result of a purely quantitative computational result may already contain some applicable information at least for the researcher. But if further transformed or “refined”, the information may become even more valuable, relevant and linkable to reality, which can be fully used to support decision making.

Linking GIS models and representations with reality is a complex and challenging process. The essence of this kind of critical approach can be summarized in one question: how to seamlessly connect abstract variables and models to correspond reality and the phenomena and semantics of real world?

2.2 GIS optimization model

The essential purpose of a GIS optimization model is to produce relevant information for a decision making process. Malczewski (1999) has made a distinction between different types of data used in spatial decision making processes. *Hard information* is described as objective data, and it consists of reported facts or quantitative data, such as remote-sensing data. *Soft information*, or subjective data, represents preferences, priorities or judgments. Malczewski argues that the way in which hard information and soft information are combined and balanced, is very central to spatial decision making.

Malczewski (1999) describes optimization simply as a normative approach to identify the best solution for a given decision-based problem. He sees an optimization method as a modeling method that searches the best (maximum or minimum) possible solution to a management problem. The problem must be structured in such way that allows the optimization to utilize it. Malczewski argues that all optimization models seek to maximize or minimize a certain quantity or quantities, which act as the objective or criterion function. Additionally, optimization problems typically have a set of constraints imposed on the decision variables. The constraints usually determine the set of acceptable solutions. Malczewski takes an example of a plant location problem, where the objective is to minimize the total transportation costs. In this case, there might be capacity constraints imposed on the transportation routes. The values of decision variables are determined by the solution to the optimization problem. If the optimization problem has more than one criterion function, it is referred to as a *multicriteria decision problem*.

Another method related to computational methods in environmental modeling is simulation. Malczewski (1999) defines simulation as a methodology for conducting experiments using a model based on the real-world system. The main difference between optimization and simulation is in the first steps of the procedures. Optimization begins with defining the objectives and actions that will fulfill the objectives in the optimum solution, whereas simulation modeling begins with defining the actions and then starts studying their effects on the objectives. The testing in simulation is done by trying out different policies under multiple external conditions. Malczewski concludes this definition by calling simulation as “the exploratory approach to decision problems” which “either reproduces a process or obtains a sample of many possible outcomes”.

2.2.1 Path optimization

Antikainen (2008) sees spatial path planning as a mean to determine an optimal path between locations. According to Longley et al. (2005), optimality in path planning context can be defined either simply as the minimal physical distance between defined locations or as the function of the distance itself, for example risk, travel time, effort or any other measure of impedance. The purpose of path optimization is essentially to minimize this impedance.

Most of the GIS tools available, including path-finding methods, are based on data models and theories that have been developed in multiple different disciplines, such as mathematics, computer science and engineering and only afterwards adopted to geography (Wright et al. 1997). According to Mainguenaud (1995), the impedance between locations and their connections can be modeled with the concepts of *graph theory*, which can be considered as a framework for abstract modeling of networks. Chou (1997) has put the graph theory in road network context as he describes it to be consisting of nodes which can be regarded as intersections of weighted edges, which are in turn considered as road segments that connect the nodes or intersections. Chou states that edges or road segments need to be assigned with a positive weight or a *cost*, which is based on the measure of impedance. After this operation, an algorithm, such as the prominent and well-known Dijkstra’s algorithm (Dijkstra 1959) may be used to find the most optimal path with a minimal impedance between different locations (Cormen et al. 1997).

Antikainen (2008) further states that the main challenge in terrain path planning is the fact that usually there are no linear infrastructure comparable to roads in the natural environment. He suggests that the only way to properly utilize the concept of path impedance is to create a solution where the natural environment is represented by the same logic in which the weighted graph model is based on. Antikainen emphasizes that while the creation of graph model represents the linkage between path planning and theory, the weighting of the edges, or road segments, is essential for connecting the mathematical abstraction with the real world.

The weights of edges or road segments or even the terrain data used, have mostly been hypothetical in multiple studies simulating path-finding issues (Antikainen 2008). However, in the case of a real world case study, the weights need to be carefully determined on a certain criteria. Antikainen recognizes the sensitivity for subjective conclusions that is related to criteria assessment, but the weights that represent traversability of a terrain can be assumed to follow the principles of least action. This concept is introduced by Burgess & Darken (2004), who have studied the *principle of least action*, which is based on the minimum effort solution. This can be defined as the tendency of nature and humans to always search for the most optimal, least action and least effort scenario or solution. In some path optimization studies, such as Atkinson et al. (2005) and Collischonn & Pilar (2000), various different weighting scenarios were used in order to compensate the lack of accurate figures.

In his study, Antikainen (2008) suggests that the main issue with connecting the mathematical abstraction with the real world is to successfully quantify the tendency for minimum effort in order to find relevant weights. In his study, he ranked different terrain types according to their attractiveness for travel. This ranking was transformed into weights that correspond to estimated movement costs. Antikainen noticed that raster data structure seems to be the more dynamic way to transform the cost estimations into a weighted graph, because every single raster cell or cell center can be transformed into a node. These nodes are surrounded by neighboring cells, and the nodes can be connected to a certain amount of these cells. These operations form the basis of the array of methods that are used in *cost surface analysis*.

2.2.2 Least-cost path analysis

In GIS environment, a path optimization problem can be approached with *least-cost path analysis* (LCP). Computation of a least-cost path from one or multiple starting points to a destination point is an interesting application in GIS. It is a process of calculating fastest, shortest or otherwise least-cost or least-impact paths over a surface of varying frictions and impediments to movement (Douglas 1994). Least-cost path analysis has been applied to various different planning purposes. Yu et al. (2003) developed a method for roadway planning by using anisotropic accumulated cost surfaces which also took possible bridges and tunnels into account. Collischonn & Pilar (2000) presented their own algorithm to find the optimal path for linear features such as roads or canals based on topography and a function relating slope, distance and cost. Bagli et al. (2011) integrated multi-criteria evaluation and least-cost path analysis to search for the most suitable route for a high-voltage power line. Choi et al. (2009) developed a raster-based GIS model that combines multi-criteria evaluation and least-cost path analysis to determine the optimal haulage routes of dump trucks in mining. Rees (2004) created a least-cost path model based on Dijkstra's algorithm and Digital Elevation Model and compared it to existing footpaths in a mountainous area of Wales. Goncalves (2010) has presented a new extension of raster-based least-cost path modeling by conducting the computation with a fixed width which is larger than one cell to produce so called wide paths. Additionally, Lee & Stucky (1998) have applied LCP to study scenic paths using visibility as a cost factor and Atkinson et al. (2005) introduced a concept for arctic road planning.

According to Douglas (1994), the process of finding a least-cost or a minimum cost path has two steps. First step is the creation of an *accumulated cost surface* from a *cost-of-passage surface*. Second step is tracing a slope line down the *accumulated cost surface* from a single departure point or multiple departure points to a destination.

Douglas defines the cost of passage surface as a dasymetric map illustrating various costs or frictions. The term "cost" is not necessarily indicating a cost directly in terms of money, but rather as a figure of resistance or impedance to movement. In a cost-of-passage surface, a higher cost indicates that an area has a high level of resistance or impedance to movement and lower costs indicate more suitable areas for movement. Collischonn & Pilar (2000) have a similar graph theory-related approach as Chou (1997) as they define the cost-of-passage surface as a grid in which the cell values are used as weights to calculate the least-cost paths. The weights correspond to the difficulty in crossing a cell.

The creation of an accumulated cost surface from a cost-of-passage surface is based on a spreading function that begins from the ending point and starts moving backwards step by step towards the starting point. An algorithm needs to be used to search the neighboring points for the point with the smallest value, in other words the least-cost point. As the accumulated cost surface continues to expand, information of the selected points are saved as the accumulated cost of the cell (Douglas 1994). The algorithm repeats this procedure until a cumulative cost value has been assigned to every cell of the grid. The accumulated cost surface is created in such manner that considers all relevant cost factors (Yu et al. 2003). After this, it is possible to trace the pathway with the lowest score line, which is generally regarded as the optimal pathway representing the minimal accumulated friction to movement from a starting point to a destination point (Lee & Stucky 1998).

According to Xu & Lathrop (1994), there are multiple different definitions of neighborhood types, such as Rook's pattern which consists of 4 cells, Queen's pattern which has 8 cells and 16-cell Knight's pattern (Figure 2). Yu et al. (2003) were able to cut off half of the cost of a path by using Knight's pattern to calculate a least-cost path compared to calculating with Queens' pattern. Also, on a general level, and especially on areas with rugged terrain, Knight's pattern did produce paths with lower total costs than the paths that were produced with Queen's pattern. In their research, Yu et al. (2003) developed a custom-made algorithm for their study purposes. In the study, it was stated that Knight's pattern was more effective in finding smaller neighboring slopes than Queen's pattern was. In this case, Knight's pattern calculation was following the contour lines and thus being able to minimize changes in slope. Yu et al. came to the conclusion that by using Knight's pattern it is possible to improve the accuracy of least-cost paths, especially in areas with steep slopes.

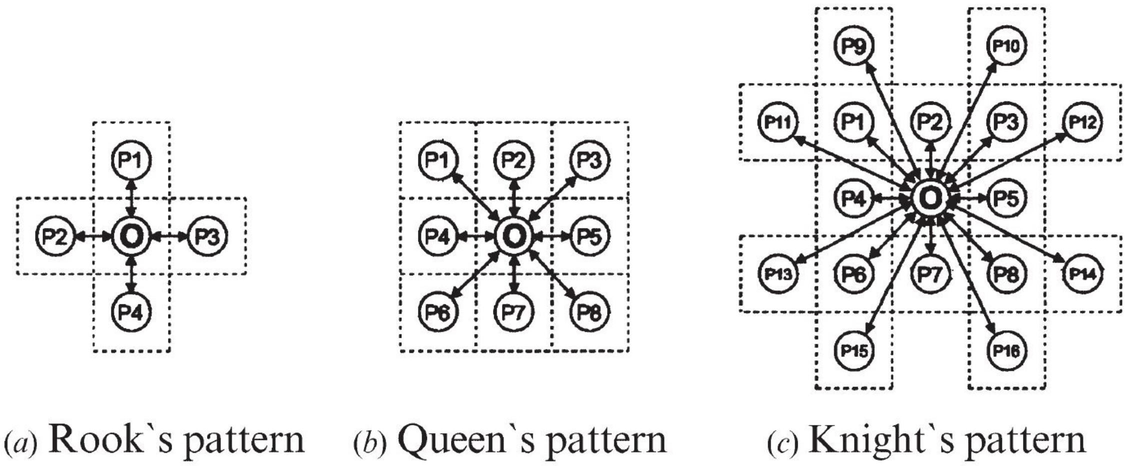


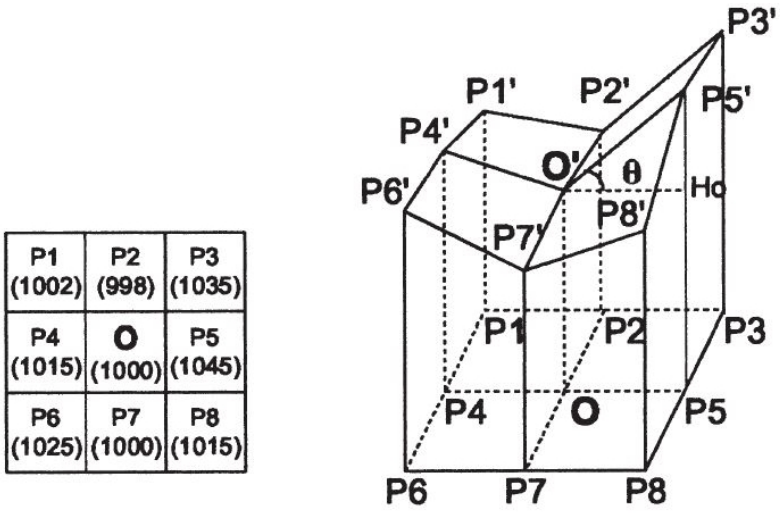
Figure 2. Three types of neighboring patterns in raster data formats (Yu et al. 2003).

Due to the dynamic nature of the simulation process of outward movement from a source, Xu and Lathrop (1995) have preferred to refer to the least-cost path analysis as *spread simulation*. The simulation is conducted on a surface or a background media. The background media or surface through which the algorithm moves can be spatially heterogeneous or homogeneous. If the media is spatially heterogeneous, there is variation in friction costs across the media and if it is homogeneous, there is no variation in friction. The media can also be categorized into two different types of simulation spread considering their properties regarding movement direction. According to Xu and Lathrop (1995), a simulation spread can be either isotropic or anisotropic. The main difference between these two spread types is that if the media either has or does not have a directional component. If the media has a directional component, a spread is anisotropic and it prefers to move in certain directions. If the directional component is ignored, a spread is isotropic and the cost of crossing a single cell is equal and non-dependent of the crossing direction, so there is no effect on whether the algorithm is moving uphill, downhill or parallel to a slope.

Yu et al. (2003) argue that in reality, there is rarely a case where isotropic accumulated cost surfaces exist, because for example in an area characterized with complex terrain surfaces, slopes in different directions are not constant. Yu et al. claim that these directional differences are important especially for roadway planning and therefore anisotropic accumulated cost surfaces should be used. In their study, Yu et al. modeled least-cost paths by using both isotropic and anisotropic simulation spread. In the anisotropic version a custom-made algorithm was developed and used for their study purposes.

Gietl et al. (2008) have reported that at the time of their study, the capability to model anisotropic costs was only a recent feature of GIS. Van Leusen (2002) argues in his archeological study that if a least-cost path analysis is based on a non-directional isotropic slope calculation derived from a digital elevation model, it has only a minor effect on the quality of the model. Another archeologist, Herzog (2014), has stated that multiple authors of archeological least-cost path studies share this view. She also suggests that on a model based on isotropic slope cost model, switchbacks or heavily undulating roads are not a favorable option, because moving parallel to a steep slope is as costly as it is for the model to move perpendicularly towards the slope. Herzog concludes that because of this, isotropic cost models never produce so-called hairpin curves. One example of an isotropic least-cost path approach has been made by Jaga et al. (1993), who used a raster grid representing slope angles in the direction of steepest descent as an attribute. They solved their road path planning problem by applying very high friction values for steep slope areas. Collischonn & Pilar (2010) suggested in their study that directional slopes are important in the planning of highways, where a zigzagging and long but less steep path is a better solution than a shorter and steeper path. They state that slope, used merely to describe the angle in the steepest direction, is very much different from slope with a directional attribute. For example, if a path crosses a slope transversely, even a rather steep slope may result in a relatively conservative gradient.

According to Yu et al. (2003), on an anisotropic accumulated cost surface, the costs from one cell to another is referred to as the spatial distance, which is the distance between two cell centers. Figure 3a shows a digital elevation model and its geometric representation. Figure 3b illustrates that the spatial distance and cell width are not equal by default, because the spatial distance is a variable dependent on changing direction and slope. For example the spatial distance from O’ to P5’ is greater than the spatial distance from O’ to P4’ because the slope angle is larger. Additionally, the spatial distance between O and P3 is greater than the spatial distance between O and P5 because of diagonal movement. Yu et al. proposes, that using spatial distance instead of two-dimensional distance will improve the accuracy of least-cost path models, because slopes between cells could lead to an accumulation of a notable distance error.



(*a*) (*b*)

Figure 3. A digital elevation model (a) and its geometric representation (b) (Yu et al. 2003).

Dijkstra’s algorithm is arguably one of the most popular algorithms used for finding last-cost paths (Xu & Lathrop 1994; Collischonn & Pilar 2000; Yu et al. 2003; Rees 2004; Antikainen 2013). According to Yu et al. (2003), the algorithm was designed to find the shortest path in a network with nodes connected by weighted links. Xu & Lathrop (1994) state that Dijkstra’s algorithm is used to determine the cumulative cost or time when a starting point, a set of sources and the costs are known. The cell centers act as the nodes of the network. One popular way to build a network for the analysis is to connect a certain amount of neighboring cells in a grid into a network. After a network is prepared, Dijkstra’s algorithm can be applied to create an accumulated cost surface, which represents the sum of weighted links between successive pairs of cells (Yu et al. 2003). After this, the algorithm examines all adjacent nodes to the node which has been most recently evaluated. The algorithm starts evaluating the nodes nearest to the starting node, and the nodes that are further away are evaluated last. The algorithm continues to work until the ending point is reached (Dijkstra 1959).

According to Antikainen (2013), the least-cost path process does also have drawbacks, which according to him, are often neglected. Antikainen argues that the main limitation of the procedure is that during the transformation process from the conceptual foundation of one-dimensional connectivity graph to a raster-based surface, is that the result can only be an approximation of the infinite number of possible movement options across a continuous space that the raster surface is representing. Therefore, even though the modeled raster-based least-cost path might be the optimum path in terms of the connectivity graph, the true optimum path in continuous space cannot be achieved (Figure 4A). However, Antikainen suggests that a raster surface with a relatively high resolution is the basis for providing even moderately accurate representation of the environment. Antikainen also claims that a least-cost path might not be unique because there can always be multiple least-cost paths reaching the destination with similar accumulated movement costs (Figure 4B & 4C).

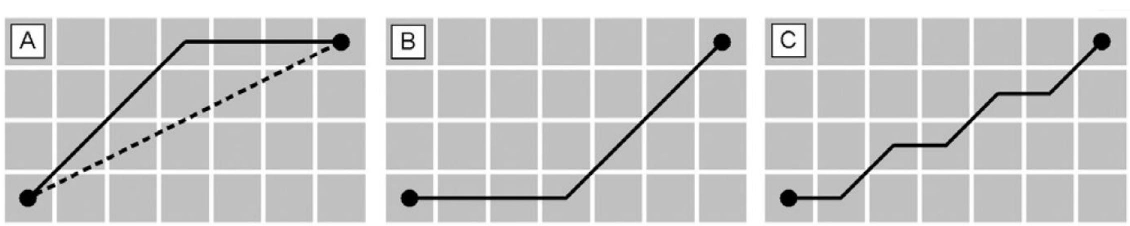


Figure 4. A homogeneous raster that allows movement between cell centers using the Queen’s pattern (3x3). A) A least-cost path (solid line) calculated between two points in

comparison to the true optimum path (dashed line) between the points. B-C) Alternative least-cost paths with equal accumulated movement costs (Antikainen 2013).

2.2.3 Multi-criteria evaluation

Atkinson et al. (2005) suggest that the cost-of-passage surface is calculated by first considering all relevant criteria that are believed to have an effect on the alignment of linear infrastructure, and then merging the criteria by *multi-criteria evaluation* or *multi-criteria decision analysis* (MCA). Goncalves (2010) has stated that the most common way to create a cost-of-passage surface or a friction surface is to use a combination of cost or friction factors. According to Collischonn & Pilar (2000), it is possible to combine several different evaluation criteria to create a grid where each cell has been assigned a value that represents the relative difficulty of moving through it.

Bagli et al. (2011) has suggested that the selected criteria should reflect the objectives that the planner desires to achieve. Usually in path optimization problems that utilize least-cost path analysis, the goal is to minimize either cost, time, risk or ecological impacts from the operation or the construction of planned infrastructure. According to Yu et al. (2003), parameters related to slope, such as slope angle and slope direction are critical factors for weighting traveling and construction costs.

Malczewski (1999) has suggested that because there are no universal techniques available for criteria selection, only the desired properties of attributes and objects may provide the guidelines for determining a set of evaluation criteria. However, the set of evaluation criteria is always problem-specific and the amount of different criteria depends on the nature of the decision problem. After establishing a set of evaluation criteria, every criterion should be separately presented as a criterion map and preferably divided into two types: *factor maps* representing the spatial distribution of an attribute reflecting the degree to which the problem objectives are achieved, and *constraint maps* representing the possible restrictions that can be used to eliminate alternatives from consideration (Malczewski 1999).

According to Atkinson et al. (2005), in a least-cost path analysis which is used to determine optimal path alternatives, the weight of different selected criteria has a direct effect on the path optimization result. Therefore the relative importance of each criterion needs to be determined in a separate process. Atkinson et al. states that this process is known as multi-criteria evaluation (MCE) and when used in GIS environment, it can be logically combined with weight development procedure called pair-wise comparison, which is part of the *analytical hierarchy process* (AHP) framework developed by Saaty (1977).

AHP is used for organizing and analyzing complex decisions. It was specifically developed to provide a consistent, quantifiable approach to problems involving multi-criteria analysis. According to Coulter et al. (2006), AHP has the potential to provide a consistent approach to the ranking of forest road investments based on multiple criteria.

As discussed, a path optimization problem consists of nodes or intersections, which are connected by edges or road segments. These edges require a positive weight or a cost to properly reflect the real world friction or impedance for movement on a surface. AHP is widely used for developing weights in GIS related problems and it also enables the combining of weights and the attribute map layers (Malczewski 1999). According to Choi et al. (2009), the decision-making process through AHP rates the importance of each criterion in relation to other criteria by using a 9-point reciprocal scale (Figure 5). If, for example there was a hypothetical problem containing factor *a* and factor *b* and the decision maker would feel that factor *a* would be strongly more important than factor *b*, the score of 5 would be inserted into an *n* x *n* matrix describing the scores (*n* being the amount of criteria). In this case, the reciprocal of that score would be 1/5, which means that *b* is strongly less important than *a*. According to Saaty (1977), the optimal set of weights can be derived from the principal eigenvector of the comparison matrix. Additionally, a *consistency ratio* (CR) is calculated to describe the consistency within comparisons. Saaty has stated that a CR of 0.100 or lower can be considered acceptable.

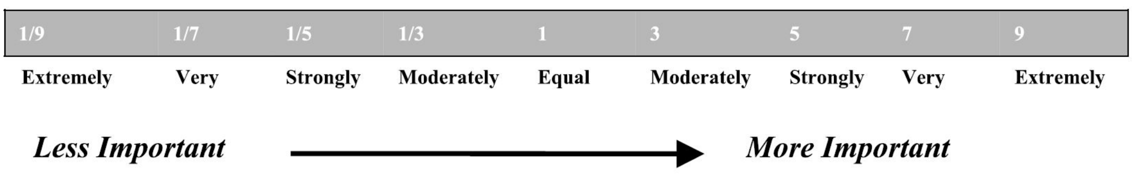


Figure 5. Nine-point reciprocal scale developed by Saaty (Atkinson et al. 2005).

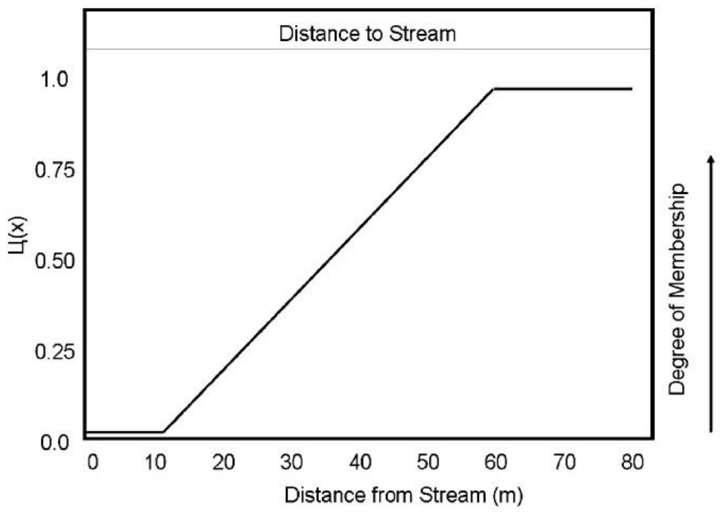
After the AHP procedure, the total priorities can be rated with *weighted linear combination* (WLC). Malczewski (2000) has described it as a technique widely applied in land use or suitability analysis for creating overlay-based aggregated composite (merged) map layers. In this procedure, the criteria will be first standardized to a common numeric range, and then combined by using weighted averaging (Jiang & Eastman 2000). According to Jiang & Eastman, MCE in GIS context has traditionally been dominated by two common procedures: the aforementioned weighted linear combination and *Boolean overlay*. In WLC, the standardization is conducted in a linear scale, for example in the range of 0–10. Boolean overlay on the other hand is based on Boolean statements, which define areas in a simple binary number system (true or false) and determine if an area is either belonging or not belonging to a designated objective. According to Eastman (1999), Boolean variables are suitable to be used as constraints, since they eliminate areas from the consideration for suitable areas. To sum up the methodology of WLC, which is the most traditional method in this context, standardized factors are first combined by multiplying each factor by a weight and then summing up the results. After this, the result might be multiplied by using Boolean constraints.

All the different factors should be standardized before presenting them in factor maps. Jiang & Eastman (2000) have identified a fundamental problem associated with conducting factor standardization in WLC. The most common method for this has been simply to rescale the values of different factors by linear transformation. Jiang & Eastman argue that the rationale for this operation is not clear. They suggest that there are many cases where it would be logical to rescale factor values to a more limited range using “anchor values”. In other words, in many instances there might be a certain threshold that when exceeded, the following exceeding values do no longer make any difference in the context of suitability. Additionally, Jiang & Eastman claim that in some cases it would seem more appropriate to use a non-linear scaling. This is based on the argumentation that because the rescaling is fundamentally conducted to express suitability, it is more sensible to define minimum and maximum levels of suitability instead of absolute minimum and maximum values during the standardization process. Eastman (1999) has suggested that these “anchor points” that are designed to correspond to the logic of set membership are superior to raw minimum and maximum values which indicate a sense of blindness to the data.

Due to these problems, Jiang & Eastman (2000) have presented a new approach in MCE called *fuzzy measures*. This approach is based on fuzzy set theory, which has been used as a bridge between human language and formal models (Zimmermann 2010). Factor standardization with fuzzy measures results in fuzzy sets, in which each class of a criterion is assigned a fuzzy value to indicate its suitability for the overall objective. In classical Boolean set theory, set membership has been determined as 1 (true) or 0 (false). Memberships in fuzzy sets are defined on a continuous, but not necessarily linear, scale from 1 (full membership) to 0 (full non-membership). As the memberships are not necessarily linear, the fuzzy values are not required to increase or decrease monotonically with class number (Bonham-Carter 1994). According to Bonham-Carter, the fuzzy values are chosen, based on subjective judgement, to indicate the degree of membership of a set.

Figure 6 presents an example of a non-linear fuzzy membership function when evaluating land suitability in relation to the proximity to a stream. The example shows that by implementing anchor values to the data (10m and 60m), the raw data values can be put under a scope that defines suitability in a more realistic manner.

Atkinson et al. (2005) have noticed, that in the context of least-cost path analysis, the polarity of fuzzy sets has to be reversed, as the objective is to minimize the cost of movement across a grid surface. Thus, as seen as a membership to a fuzzy set, values closer to 0 indicate stronger membership in the set of suitability for path.



$$Ц(x) = \begin{cases} 0 & X < 10 \\ \dfrac{X - 10}{50} & 10 < X < 60 \\ 1 & X > 60 \end{cases}$$

Figure 6. Fuzzy membership function and graph for stream data (Atkinson et al. 2005).

In the context of utilizing multi-criteria evaluation for simulating a least-cost path in GIS environment, Atkinson et al. (2005) used fuzzy measures to standardize factor scores in order to use pair-wise comparisons as part of AHP to weigh different criteria. Atkinson et al. used a fuzzy logic combination functions to combine the fuzzy sets through algebraic multiplications and sums. In the same context, Choi et al. (2009) used fuzzy membership functions to standardize factors, determined the weights for each factor with pairwise comparisons and finally used WLC to combine factor and constraint scores.

2.3 Wind power infrastructure

Onshore wind power parks require an extensive network of infrastructure. The fundamental infrastructure of wind parks is usually divided into two categories: civil works and electrical works, together referred to as the *balance of plant* (BOP). The BOP is usually designed and installed by a third party, such as an infrastructure company, separate from the project developer and turbine manufacturer (EWEA 2009: 101). Civil works include a road network, construction areas with foundations, meteorological mast and utility buildings. Electrical works comprise of electrical grid and all necessary buildings and utilities for transmitting electrical power (EWEA 2009: 100). A more simplified approach to the classification of different infrastructure elements comes from a chronological life cycle perspective: from production and transportation to installation, grid connection and finally maintenance (U.S. Department of Energy 2015).

There are different ways to classify infrastructure elements, but in this study a simplified three-element classification is used. This classification emphasizes the most essential components in the process of infrastructure alignment. Following classification is used:

1. Access roads for transportation and maintenance
2. Construction areas for foundation, construction and assembly, and erection of turbines
3. Internal electrical grid with necessary buildings and utilities for transmitting electrical power and turbine operation data as well as weather conditions

2.3.1 Access roads and transportation

The components of a wind power plant are usually transported along roads by large special vehicles. The tower is carried in multiple segments and the turbine, blades and nacelle are all transported separately (Gasch & Twele 2012: 499; STY 2015). All elements are transported to the wind power area, where the turbine will be assembled, erected and finally operated. After the erection phase, logistics infrastructure such as the access roads and wind power areas will be cleaned up and used mainly for maintenance purposes. Operation time is typically about 20 years, after which the power plant will be dismantled (Gasch & Twele 2012: 503). Subsequently, a new cycle may begin if the same sites will be used to install new, more advanced wind turbine technology. This process is referred to as *repowering* (Gasch & Twele 2012: 507).

Turbine parts cannot be transported by using just any regular roads. There are strict specifications and requirements for these roads considering road shoulder width, side and upper clearance, flatness, carrying capacity, curve angle, curve radius and slope steepness (Gasch & Twele 2012: 499). This is mainly due to the fact that a single blade of a modern onshore wind turbine can be up to 60–70 meters long. Because of the large dimensions of modern wind turbines, the driving surface of access roads generally needs to be 4 to 6 meters wide. Also, areas alongside the roads are cleared of any obstacles such as forest or large rocks, as the required clearance width and height might reach up to 6–8 meters. Additionally, the roads need to be able to support large construction cranes and fully loaded trucks, potentially weighing about 150 to 200 tons. Each wind turbine manufacturer specifies the road requirements for their own turbines (Wizelius 2015: 178) and these documents are generally classified as confidential information.

In Finland, Vaasa Energy Institute (2010) has evaluated the construction costs for a new forest road suitable for wind power plant transportation to be in the range of 200 000–300 000 euros per kilometer. However, if an existing road is even in adequate condition and it is allowed to be used for wind power plant access purposes, it only needs to be upgraded. In this case, less work is required and the costs are generally considered to be around half of the construction costs of a new road. The costs may be substantially higher if the terrain is challenging, for example swampy areas, where extra groundwork is needed to reinforce the soil. Another important cost factor is related to the requirement for maximum allowed road gradient, which is usually in the range of 8–12 percent, or 5–7 degrees. The costs are dependent on the amount of surface removal or rock demolition works required to make the surface gradient correspond to the requirements.

The total costs for access road network alignment planning and construction are generally considered to be strongly case dependent (Gasch & Twele 2012: 494). Vaasa Energy Institute (2010) has suggested that the total costs for roads for a single 2 MW onshore turbine located in Europe would be around 1 % of the total project costs. It can be assumed that for bigger projects with a larger amount of turbines, the price per turbine is smaller due to quantity discount but the total price of road planning and construction is higher due to the need for a more extensive road network. Additionally, depending on the project location, existing infrastructure such as existing roads also have an effect on the total costs.

2.3.2 Construction areas

In municipality-level planning in Finland, the site where a wind power plant will be located and operated is usually referred to as the wind power area in a master plan. The wind turbine, including the rotating blades, must not exceed the boundaries set for a wind power area. In more specific planning regarding the assembly and erection, a *construction area* will be planned inside the wind power area. Construction area, also referred to as work area, is the site around where the wind power plant is assembled, erected and within the foundation is constructed (Wizelius 2015: 178). Directly next to the foundation, this area includes predetermined locations for the lifting crane and auxiliary cranes, temporary storage areas for components and optionally housing and offices for the construction personnel. Different turbine manufacturers have different requirements for the dimensions and the specific overall layout of a construction area.

The size of a construction area may vary a lot depending on turbine manufacturer’s preferences, project developer’s preferences and project specifications. Generally, the size can be anything from 3 000 to 10 000 square meters. A larger size might minimize the project management-related risks regarding smooth assembly, erection and logistical procedures. On the other hand, extensive groundworks and forest cuts increase the economic and ecological effects. Additionally, if a wind power plant is sited in a forested area, a larger patch of clear-cut and leveled area among otherwise dense forest may affect the wind dynamics and increase the turbulence in the proximity of the wind turbine (Gasch & Twele 2012: 126).

In larger wind parks, the construction areas are usually constructed in a way which allows the access roads to go through the area. In this sense, these areas can be perceived as relatively large nodes in the access road network. During the construction phase, the construction areas have a vital role in the smooth operation of project logistics.

2.3.3 Internal electrical grid and external grid connection

When wind turbines are in operation, the generated electrical power has to be transmitted to the national *grid* so it can be distributed to the consumer. Every wind turbine generator has its own low voltage power system. The generated electrical power is first transformed inside the turbine to medium voltage and then transmitted through a medium voltage internal grid to a *connection point*, where the electrical power is then transformed to high voltage and then further transmitted to the grid point where it is distributed to the consumer. The connection point can also be referred to as *point of common coupling* (*PCC*). (Empower 2014)

The objective of the internal grid is to transmit the generated electricity from each wind turbine to the PCC. In the internal grid, medium voltage cables, around 20 kV–35 kV, are used (EWEA 2009: 102). Cabling can be carried out by using either overhead cables or underground cables, but according to Empower (2014), despite being significantly more expensive, underground cables are almost exclusively used in Finland. The cable trenches for medium voltage and high voltage lines are generally the same size, both are about 1 meter in width at ground level.

As mentioned, the electrical power is transformed to high voltage in the PCC. In practice, a point of common coupling is an *electrical substation* equipped with a transformer. In wind power projects, this kind of collector substation is usually required to "collect" the electrical power from all of the wind turbines and then transform it to a higher transmission voltage for the external grid. A substation without a transformer is referred to as a switching substation. Usually in wind power projects, there is a dedicated collector substation at the PCC, and a high voltage cable, usually 110 kV, is built from the PCC to a switching substation (grid point) owned by the grid operator. This switching substation acts as the gateway to the external grid. Berzan et al. (2011) call this network of cables and transformers that harvest the energy from wind turbines and make it available to the grid as a *collector system* (Figure 7).

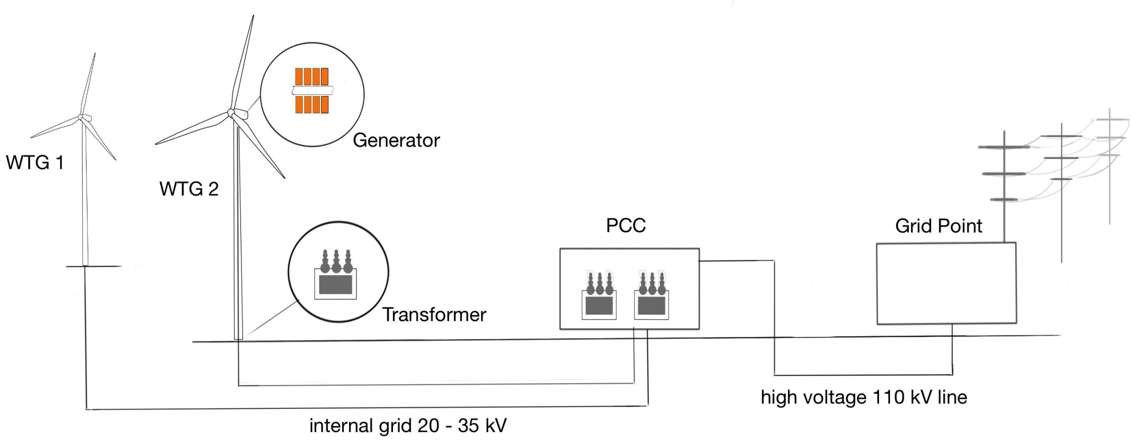


Figure 7. An example of a wind park collector system (EPO 2015a).

The requirements for a PCC vary according to the total rated power of a wind power project. In most cases, the outbound voltage of a PCC is in the range of 110 kV to 400 kV, depending on the total capacity of the project and the specifications of the PCC. Empower (2014) has made an rough estimation that if the total rated power of a wind park is between 25 MW and 200 MW, a 110 kV substation should be sufficient. For over 200 MW wind parks, a 400 kV substation is recommended.

According to Vaasa Energy Institute (2010), the total costs for electrical grid construction for a single 2 MW onshore turbine located in Europe would be around 9 % of the total project costs. The total price for a 20 kV underground cable is around 79–140 euros per meter depending on the terrain conditions and the cable specifications (Energy Department 2010). This price range includes installation works, but groundwork costs are excluded as they are terrain dependent. Similarly to access road construction, the costs of building an electrical grid for a wind park are highly case dependent. The main costs consist of cables, the groundworks and the potential substation equipment. Similarly to road construction, the cable costs correlate with overall size of a project area, involving the distance between turbines, PCC and external grid connection point.

2.4 Infrastructure alignment planning

2.4.1 General practices in infrastructure alignment planning

From the perspective of land planning, the planning scope is very similar between access roads and electrical grid. The task is essentially to create a new network and then connect it to an existing network in the most optimal way. The internal road network will be connected to the national road network and the internal grid network will be connected to the national grid network. In other words, the internal networks will be integrated into the existing external networks as effectively as possible. Thus, the locations of connection points for the roads and the electrical grid are rather deterministic factors in the infrastructure planning.

After the municipal planning phase for designating wind power areas in the master plan, the final locations of the wind power plants are defined. The process of wind park infrastructure alignment planning generally continues by planning the access road network for the wind power areas. It has been almost a normative approach in Finland, that the electrical cables are installed in ditches alongside access roads, because it is easier and also less expensive than dragging the cables through an untouched terrain (Vaasa Energy Institute 2010; Empower 2014). It can be pointed out, that cable ditches are notably narrower and also not as much affected by topography than access roads. On the other hand, it is also possible to install cables above the ground, but inside steel pipes. According to Wizelius (2015: 178), if not laid in ditches alongside access roads, reinforced steel pipes can also be used to integrate cables with the roads. In overall, grid construction is more flexible than road construction in terms of possible construction methods. However, the first established wind park projects in Finland have been relatively small, and located close to existing infrastructure. It seems that this has made it relatively simple to come up with an alignment plan, in which the external infrastructure network can be reached with a solution of one combined internal infrastructure network.

There are some general constraints for infrastructure alignment planning that apply to the alignment of both underground cables and access roads. The overall construction effort is minimized if obstacles, slopes, wet soil and water areas are avoided. Anderson & Nelson (2004) have stated that from an economic perspective forest roads can be seen as an investment in the transportation system and also as an increase in the asset value of a forest. On the other hand, Anderson & Nelson claim that forest roads consume potentially productive forest land areas and in some cases, may become environmental and economic liabilities. It can be argued that the less road construction works such as ground leveling or demolition works are carried out, the less economic and ecological impact will occur. Nevertheless, excessive evading of the aforementioned constraints may result in longer road and cable network resulting logically in higher road material and copper and aluminum costs. In this sense, there is no straightforward solution for this problem because minimizing road length nor minimizing the amount of demolition works may not result in the most optimal solution. Additionally, there might be some construction restrictions specified in the planning directives. For example, any construction work is legally forbidden in an area which has been determined as a living habitat of a protected species.

Another aspect in infrastructure alignment planning is a logistical one. During the construction phase, a large amount of trucks and other vehicles are moved inside the project area, delivering components or assisting in assembling and erecting the wind power plant. This process involves a lot of workforce and rented machinery and if prolonged, it can have an effect on the total costs and timetable of wind park construction. This emphasizes the importance of smooth and dynamic logistical operation. The risk of prolonged construction phase may be minimized by implementing a transportation management plan including defined entrance and exit points to the project area and by planning passing lanes or turnouts for trucks in logistically critical areas. Additionally, Y-shaped junctions should be planned instead of T-shaped junctions to avoid sharp curvature. Also, the same road crossing should be possible to travel in both directions, even with long vehicles.

If assumed that the road alignments are more difficult and more expensive to adjust, then the electric grid must be aligned with the access roads. In this case, the roads are generally planned before electrical grid because of the larger impact on the landscape (Petri Koski, personal communication 29.1.2015). However, the planning of the electrical grid might become more complicated if for example a couple of turbines are located behind obstacles or otherwise challenging terrain and the access road comes from the opposite direction than where the substation is located. In a case like this, aligning the cables with the road, the total length of cables might increase so dramatically, that dragging the cable straight towards the substation could still be a less expensive option, even when taking also the required extra groundworks into account. Especially in larger, over 10-turbine projects, it can be argued that the possibility of electrical cables taking “shortcuts” from the access road network should be investigated. Additionally, for longer distances, the cables can be combined in a cable cabinet or a cable hub (Figure 8). In this case, a more expensive cable with a larger diameter and higher conducting capacity is required, but all the power can be transmitted along a single path and with less losses.

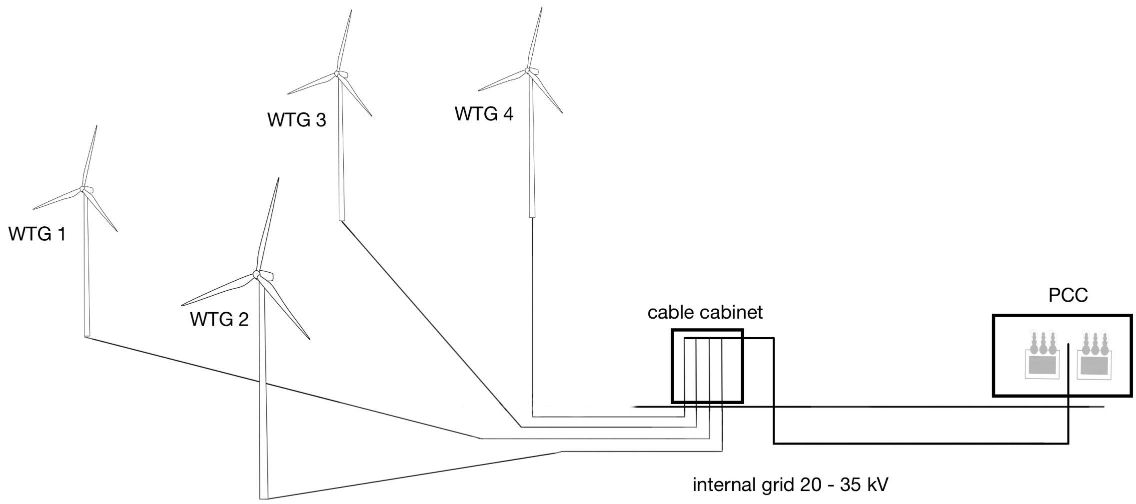


Figure 8. A four-turbine circuit connected to PCC through a cable hub (EPO 2015b).

In a more scientific method for planning of grid topology, Berzan et al. (2011) have argued that designing the collector system (Figure 9) for a wind park is a complex problem, and lower construction and operation costs can be achieved by good layout design. To simplify the problem, sets of turbines are connected to substations using cables. All the turbines that are connected to the same substation form a *tree topology*. Each turbine does not necessarily have to have a dedicated cable connecting the turbine with the substation. A set of turbines can be connected to each other and only one is connected to the substation. In this case, every segment of cable is required to have enough transmission capacity for the number of turbines downstream. Every branch that is rooted at the substation is called a *circuit* (Berzan et al. 2011). Berzan et al. conclude, that the amount and locations of substations and electrical cables have the largest effect on costs. To minimize the costs, the location of substations should be optimized according to the locations of turbines and the grid connection point. Some project developers may opt to locate the substations as close to the grid connection point as possible, as the licensing process for high voltage cables can be challenging. Also, as mentioned, high voltage cables are much more expensive than medium voltage cables. Additionally, the cost of underground cable construction depends on the terrain it is installed in. The type of substations and cables can vary a lot and these variables have to be evaluated case by case.

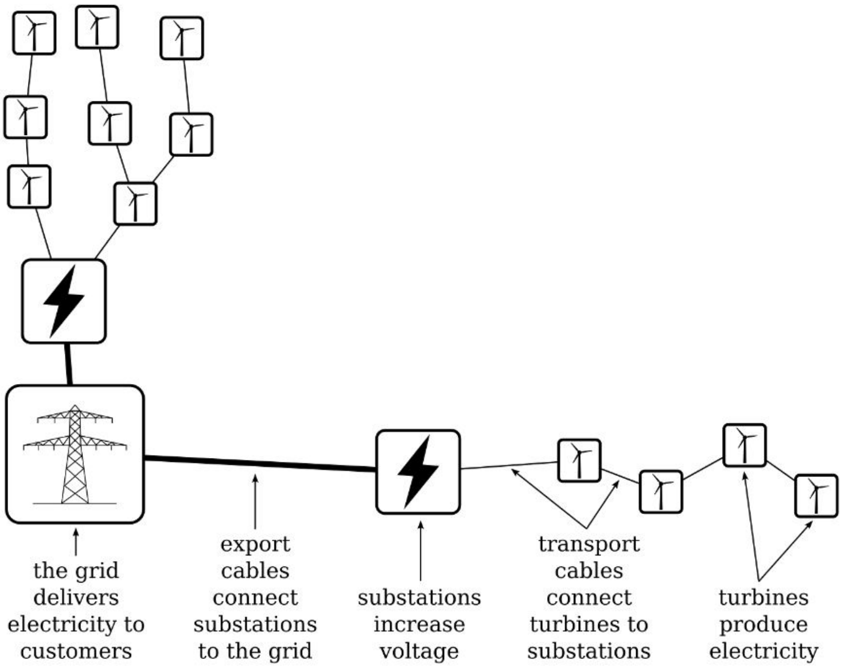


Figure 9. An example layout of a wind turbine collector system (Berzan et al. 2011).

2.4.2 GIS in infrastructure alignment planning

According to Wizelius (2015: 182), there are various GIS-based tools available for the planning of wind power projects. However, these software generally emphasize the siting of wind power plants and wind measurement-based calculations such as energy output and turbulences (Wizelius 2015: 182–183). The planning of infrastructure alignment for wind parks is usually carried out in a similar way than the planning of any other infrastructure. Access roads are designed by construction companies, and electrical grid topology is designed by electric power infrastructure companies. For example, the access roads are well within normal civil engineering practices (EWEA 2009: 102) and the configuration of internal electrical grid is conducted by an electric power engineer (Wizelius 2015: 178). However, planning the alignment of infrastructure in areas with dense forest or varying topography can be a challenging and complex process, which makes it an interesting task also from a geographical aspect. According to Wizelius (2015: 178), depending on the character of the terrain, the construction of a wind turbine access road network can either be "simple" or "very complicated". Also, the planning process of a forest road network can be difficult and time-consuming. While some might use planning software such as AutoCAD, a large portion of civil engineers are still using different thematic paper maps, such as contour and terrain maps describing topography and land cover classification (Liu & Sessions 1993; Rogers 2005; Picard et al. 2006). Anderson & Nelson (2004) describe this type of manual planning as labor-intensive and time consuming. By utilizing tools and practices offered by Geographic Information Systems (GIS), it is possible to save days or weeks of field work and more importantly, come up with a more sophisticated design than with traditional practices (Rogers 2001; 2005).

Compared to traditional land use planning such as road planning, which is very much based on practical experience and knowledge, GIS offers an information system based approach which can be applied to land use planning. According to Goodchild (2009), GIS can improve this task by allowing the processing, analysis, modeling and storage of geographical data and information. Analysis has often been perceived as the key component of GIS through searching of patterns and anomalies from geographic data, but as GIS is an integrated technology, it can also be easily incorporated with data acquisition through *remote sensing* (Malczewski 1999; Goodchild 2009). In land use planning such as road planning, the geographical data describing elevation, detailed topography, vegetation, land cover patterns and existing roads is absolutely essential to achieving a good design. Incorporating remote sensing data with GIS analysis tools can produce more accurate and even new types of information about the planning site than a field specialist who observes the site and takes photographs and makes notes. In addition, it can be argued that remote sensing data is less sensitive to subjective observations and conclusions.

Remote sensing is used for obtaining information about the Earth through a device that is not in a direct contact with the studied phenomenon (Lillesand et al. 2008). Usually remote sensing is carried out using satellites or airborne vehicles equipped with sensor devices. According to Lillesand et al., remote sensing is helping us to better understand the earth as a system. One technique that is used for accurate mapping of terrain elevations is *Lidar* (*light detection and ranging*). It is an active remote sensing technique, which transmits pulses of laser light towards the ground and then measures the time it takes for the pulse to return (Lillesand et al. 2008). Since the late 1970s when lidar was first used for topography mapping, the technique has seen major developments. Lillesand et al. suggest that nowadays modern lidar-based data acquisition involves dedicated aircrafts for photogrammetry equipped with reliable and accurate GPS, measuring and computer devices, robust data storage, fast laser, and better overall cost-efficiency. Lillesand et al.

2008: 627) state that the outlook for applications of lidar systems is "an extremely promising one".

2.5 Process Modeling

In the context of process modeling, models are perceived as a method to explain, describe and document a collection of activities. According to Bannon (1995: 67), "Models are thus seen, in our view, as interpretations, as constructions, which for some purposes, under certain conditions, used by certain people, in certain situations may be found useful, not true or false."

Process modeling has emerged from Business process management (BPM), which in turn has its roots in information system sciences, where information technology and operations management meet. It is a concept based on the recognition that every product that a company creates to the market is the outcome of performing multiple different activities (Weske 2012). According to Weske, these activities are instrumental in managing and organizing these activities in order to better understand their interrelationships. Additionally, business process management is getting more and more supported by the utilization of information technology and information systems. Weske has noticed, that it is a common situation in many companies that the organizational business aspects do not always correspond to the information technology that companies are using. In addition to being an essential component in the understanding of how companies operate, business processes can also be crucial in designing and realizing flexible information systems.

If business processes are described graphically, it is called *process modeling*. Weske (2012) suggests that by ordering different activities in a process, a company can use the outcome as a blueprint for organizing work activities. Weske states that a business process model which consists of a set of activity models and execution constraints between them acts as a blueprint for a set of business process instances. Following this logic, it is easy to repeat similar process instances by using the same process model.

2.5.1 Business Process Management Notation

According to Weske (2012), there are various different graphical notations for process modeling but they are essentially very similar to each other. Business Process Model and Notation (BPMN) is probably the most commonly used graphical notation for modeling and documenting business processes in a holistic and standardized way. It has been developed to be understandable for all stakeholders or such, regardless of branch of science. Smith & Fingar (2003) argue that because business people have an essential need to be able to communicate in a way that they find comfortable, any standard process management language cannot realize its full potential until also the business-oriented people can use it to communicate.

According to Smith & Fingar (2003), BPMN is the graphical visual notation of the Business Process Modeling Language (BPML). BPML is a XML-based specification for building process management systems and for modeling business processes and it provides the abstract model for all processes. Smith & Fingar claims that while BPML is the technical language and tool for process management used by technical users such as IT experts, BPMN uses a set of drawing symbols to represent BPML elements so that non-technical users in an organization, such as business people may fluently understand and communicate with the technical staff. Additionally, despite the fact that there is a one-to-one correspondence between BPML (code) and BPMN (diagram), the main objective of BPMN is to meet the need for process communication.

Weske (2012) illustrates that in BPMN, several diagram types for describing both process choreographies and process orchestrations are defined. According to Weske, the term choreography comes from the fact that there is no central agent to control the activities. In addition, the situation is comparable to a dance show, where a certain mutually agreed choreography takes place and all the dancers work autonomously. In a BPMN diagram, organizations are represented by “pools” and “swimlanes”. Pools contain a specific organization and lanes contain organizational entities such as departments in organizations. BPMN also consists of activities, tasks, events, sequence flows and gateways. For a full BPMN manual, see Weske (2012).

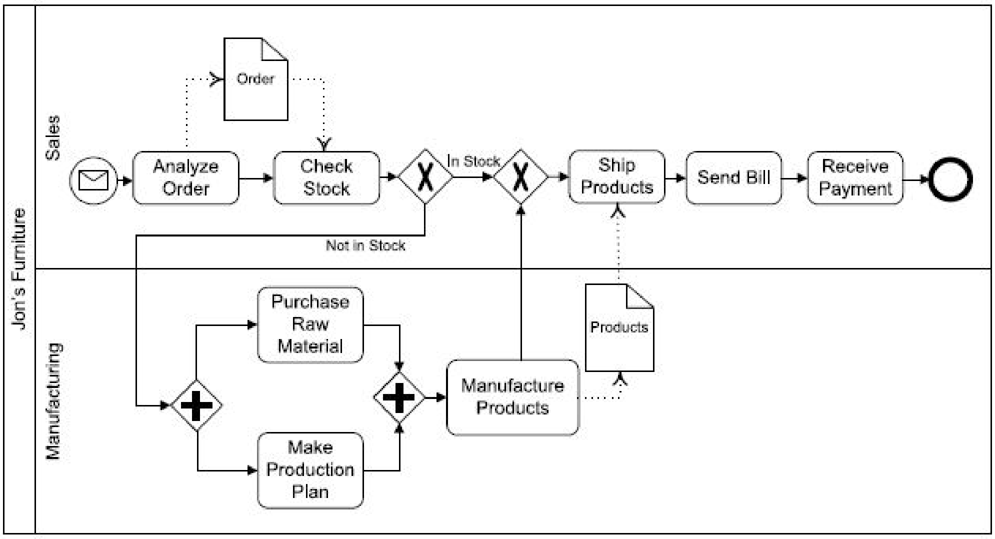


Figure 10. An example of a business process diagram with role information in BPMN (Weske 2012).

2.5.2 Metaplanning

Land use planning can be a complex process and it may often involve various steps, procedures which might also be iterative. These steps and procedures often tend to be conducted by multiple actors with different scientific backgrounds. There might also be various different organizations involved in a single planning process. Campagna et al. (2014) have conducted pioneer research on incorporating metaplanning with process modeling. They define metaplanning as the design of the planning process. The need for metaplanning arises from the growing need to manage complex multi-actor planning processes and procedures, which usually involve difficulties in collaboration and lack of mutual understanding among the various different actors. Campagna et al. argue that metaplanning can help actors such as internal or external stakeholders or observers to understand how and when planning decisions are made, why they are being made and by whom. The concept of metaplanning aims to specify activities, actors, tools, methods, input and output data, and workflows involved in the planning process.

Metaplanning ultimately utilizes the concepts and tools of BPM. According to Campagna et al. (2014), a process-oriented way of thinking has been increasingly emphasized during the last decade and simultaneously BPM techniques and tools have experienced substantial developments in order to improve process management. Therefore Campagna et al.

suggest that BPM could act as the methodological and technical approach for implementing metaplanning to support the process improvement and simplify the integration of external information systems. In the context of improvement and development of processes, the planning-related process models can either be descriptive (as-is) or prescriptive (to-be). While BPM was originally developed for describing and managing business processes, Campagna et al. feel that BPM techniques can be applied in in land use planning as well for developing the explicit design of a planning process.

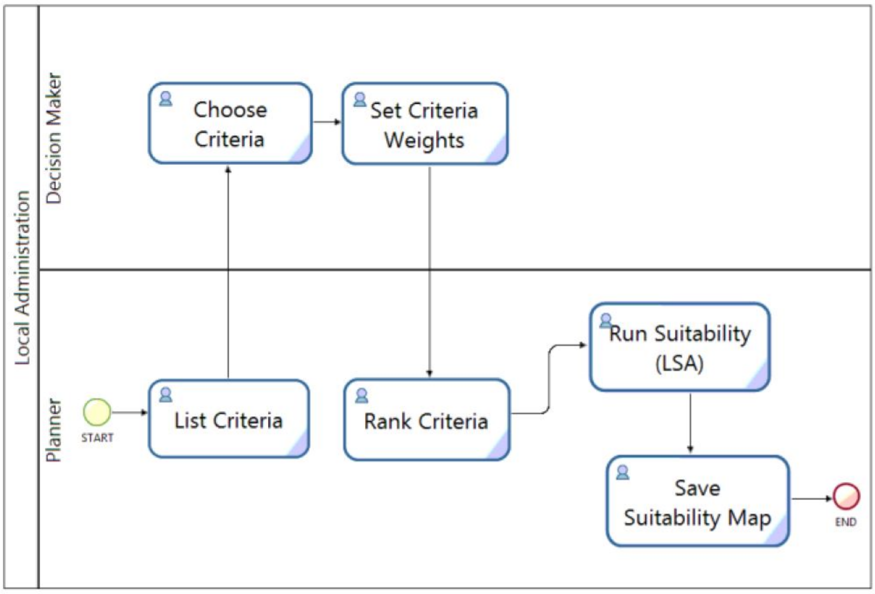


Figure 11. Land suitability analysis business process diagram in BPMN (Campagna et al. 2014).

3 STUDY AREA

3.1 Nordanå-Lövböle wind park project

Nordanå-Lövböle wind park project is located in Southwestern Finland on the island of Kemiönsaari (Figure 12). The project is owned and developed by Varsinais-Suomen Energia Oy / Egentliga Finlands Energi Oy (EFE). The master plan for the wind park includes 29 wind power plants in the area of approximately 1600 ha. The plan describes designated wind power areas for each of the 29 wind power plants. The planned wind power plants will be 100–145 meters high at the hub height and depending on the wind turbine model, the blades may reach up to 205 meters in total height. The nominal power output for each turbine will be 2–4 MW. The master plan (Kemiönsaari 2014) for the project was approved by Kemiönsaari council in their meeting on 9th of December 2014. (EFE 2015a)

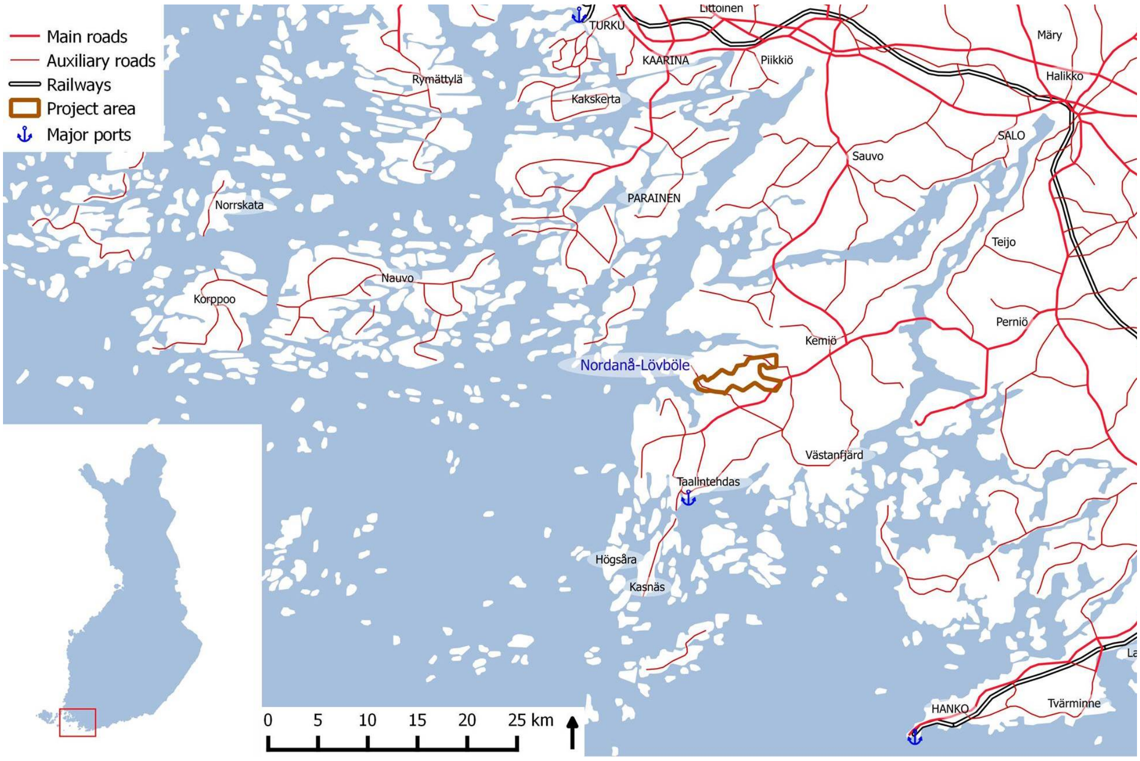


Figure 12. The 1600 ha large Nordanå-Lövböle wind park project is located 70 km from Turku and 100 km from Hanko by road. (EFE 2015b)

Land use in Nordanå-Lövböle is dominated by forestry and mining activities. The topography is varying and undulating from sea level up to 70 meters as shown in Figure 13. Nordanå-Lövböle is a suitable site for wind energy production for several reasons. First and foremost, the wind conditions in the area are excellent. Secondly, because of the forestry and mining activities, the area has a predominantly industrialized character. Thirdly, from the perspective of infrastructure planning, there is an electrical switching substation nearby the project area and the national road network passes by the project area. The area is designated as a potential wind energy production site in the wind power study conducted by the Regional Council of Southwest Finland in 2011 (Klap et al. 2011) and marked as such in the Regional plan for wind power. According to a vast array of studies conducted during the environmental impacts assessment (EIA), there are no official natural, cultural or landscape-related protection areas in the project area. However, there are few small archeological artifact spots and protected animal breeding grounds in the project area which are recommended to be avoided.

The land on the project area is owned by private land owners and EFE has effective land lease agreements with each of them. Planning-wise the land lease agreements do not contain any specific restrictions for infrastructure planning, such as roads, construction areas and electric cables. However, there are agricultural activities in the area and it would be desirable to avoid construction works in these areas. On the other hand, new roads are generally welcomed by forest owners, because the roads increase the value of the land as they are an important element in forestry logistics.

The project area is relatively large, which means that the road network will be quite extensive as well. A total of about 24 kilometers of internal road network has been initially planned by using manual methods based on shortest distance calculations. This road network was presented in the master plan. Additionally, at the time of this study, EFE had received some undisclosed offers for the construction of internal grid from companies that carry out engineering, procurement and construction services (EPC). Even though the offers were very preliminary, some figures for the total grid length were already estimated. According to the preliminary single line diagrams EFE had received, the total length of medium-voltage cable network was estimated at around 34 kilometers. A smaller cable type connecting turbine circuits with each other accounts for 15 kilometers and a larger cable type connecting a single turbine to the substation accounts for 19 kilometers.

Nordanå-Lövböle wind power project is a suitable case for testing the methodology presented in this study for infrastructure alignment planning. Firstly, it is a large wind power project on the Finnish scale. At the moment, there are not any operational wind parks in Finland of the same size. Once in operation (estimated 2018), it will be the biggest wind park in Southwest Finland. Secondly, the varying topography of the area poses challenges for infrastructure construction operations. The mainly forested project area does have some private roads, but these roads are generally narrow and not in a condition to support heavy construction machinery. These characteristics make the planning of infrastructure alignment in the area a challenging and multi-dimensional task.

There are also differences within the project area. The planned wind turbines located on the western side of the project, are sited on a more flat terrain and on average, closer to a road than the wind turbines on the eastern side of the project. On the other hand, the distance from the main substation is significantly longer to the turbines on the western side than to the turbines on the eastern side.

Due to the topographical profile of Nordanå-Lövböle area, it might have higher total infrastructure costs than a typical realized wind power project in Finland. Extensive road network built on an uneven, rocky terrain might notably increase the relative costs of road infrastructure. However, this is the growing trend in wind power planning in Finland and also in other countries: as the easiest and the most suitable sites are generally realized first, the search for new, possibly more challenging sites begins. This trend creates pressure on the planning of more cost-efficient layouts for wind park infrastructure especially in planning-wise more challenging areas.

In Figure 13, the main infrastructure planning elements of the study area are overlaid on a Digital Terrain Model (DTM) owned by EFE Ab. Locations for cable hubs and cabinets, excluding the main substation, have been derived from an unpublished preliminary plan produced by EFE Ab. The data for wind power plant locations, main substation location and project are borders were derived from the master plan (Kemiönsaari 2014). Existing roads were exported from the nationwide *Digiroad* dataset produced by Finnish Transport Agency (FTA 2016).

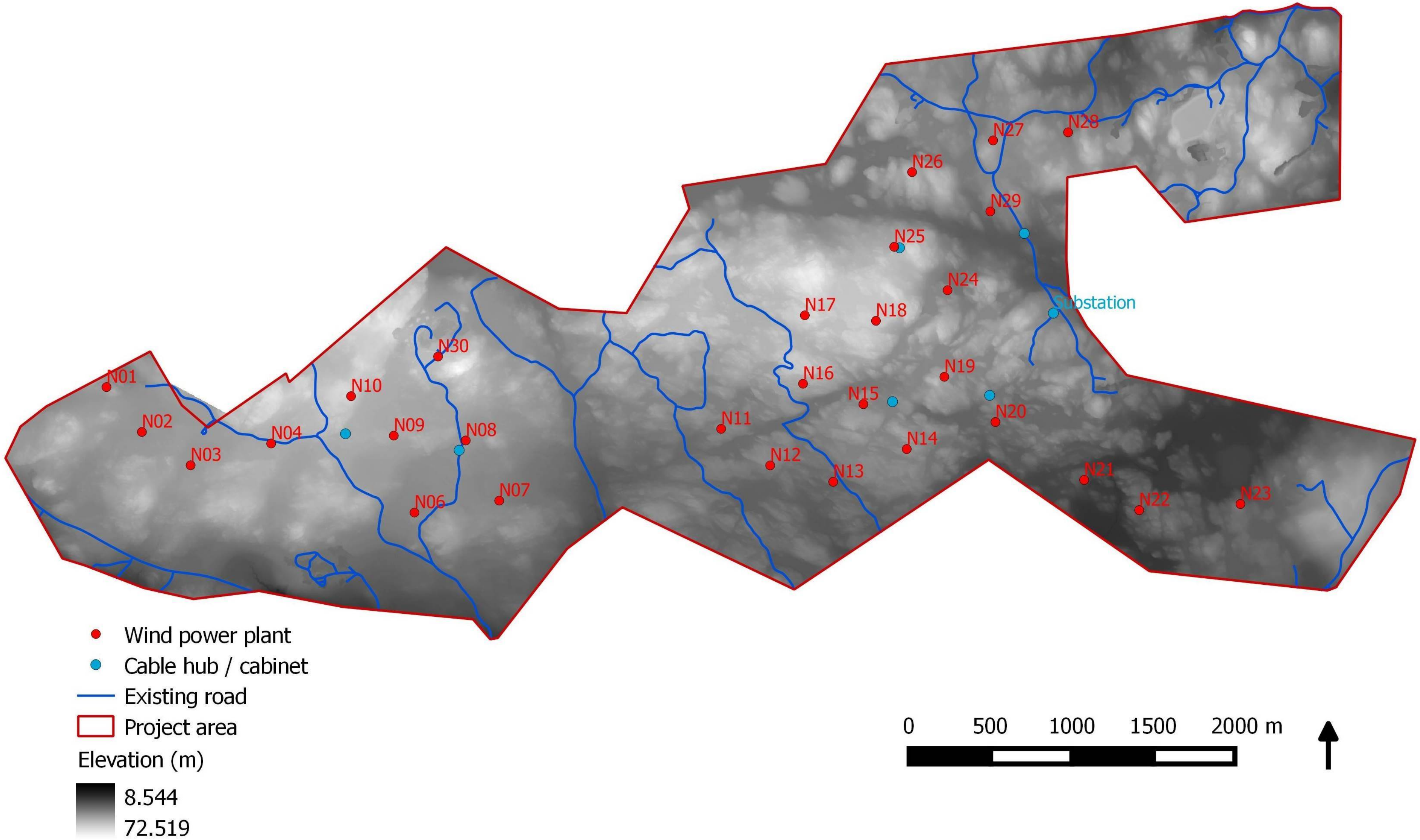


Figure 13. The elevation in the project area varies over 60 meters, from 8.5 to 72.5 meters. Digital Terrain Model (DTM) provided by EFE Ab.

4 MATERIAL AND METHODS

This study was based on an optimization model where real-world remote sensing data and planning restrictions data were loaded into an optimization algorithm through geospatial data processing. The purpose of the optimization algorithm is to simulate realistic minimum effort movement on a surface while simultaneously complying with the planning restrictions. The remote sensing data was re-classed for criteria evaluation. This resulted in factor- and constraint maps describing the degree of suitability for paths. These maps acted as the source data for the path optimization algorithm with the maps being weighted and combined in various ways for multiple optimization scenarios with different objectives. Finally, the whole process was documented by using methods developed for process modeling.

In Figure 14, the overall research strategy and working steps are presented. The flow diagram is vertically divided in three sections describing the workflow with three different approaches: practical, technical and methodological. Although different methods were used in different phases of the study, Geographical Information System (GIS) was acting as the base platform for editing, storing and presenting the research data.

Almost every step of the least-cost path analysis was conducted with open source software to allow maximal repeatability of the analysis. Only the conversion of lidar data to raster grid models was done by using a commercial software, ArcMap 10.1, to ensure the best possible quality of conversion. Successfully running the analyses completely with free software is a strong indication of the recent development of available open source GIS software. QGIS 2.8.2 and GRASS GIS 6.4.3 were the primary software packages used in this study.

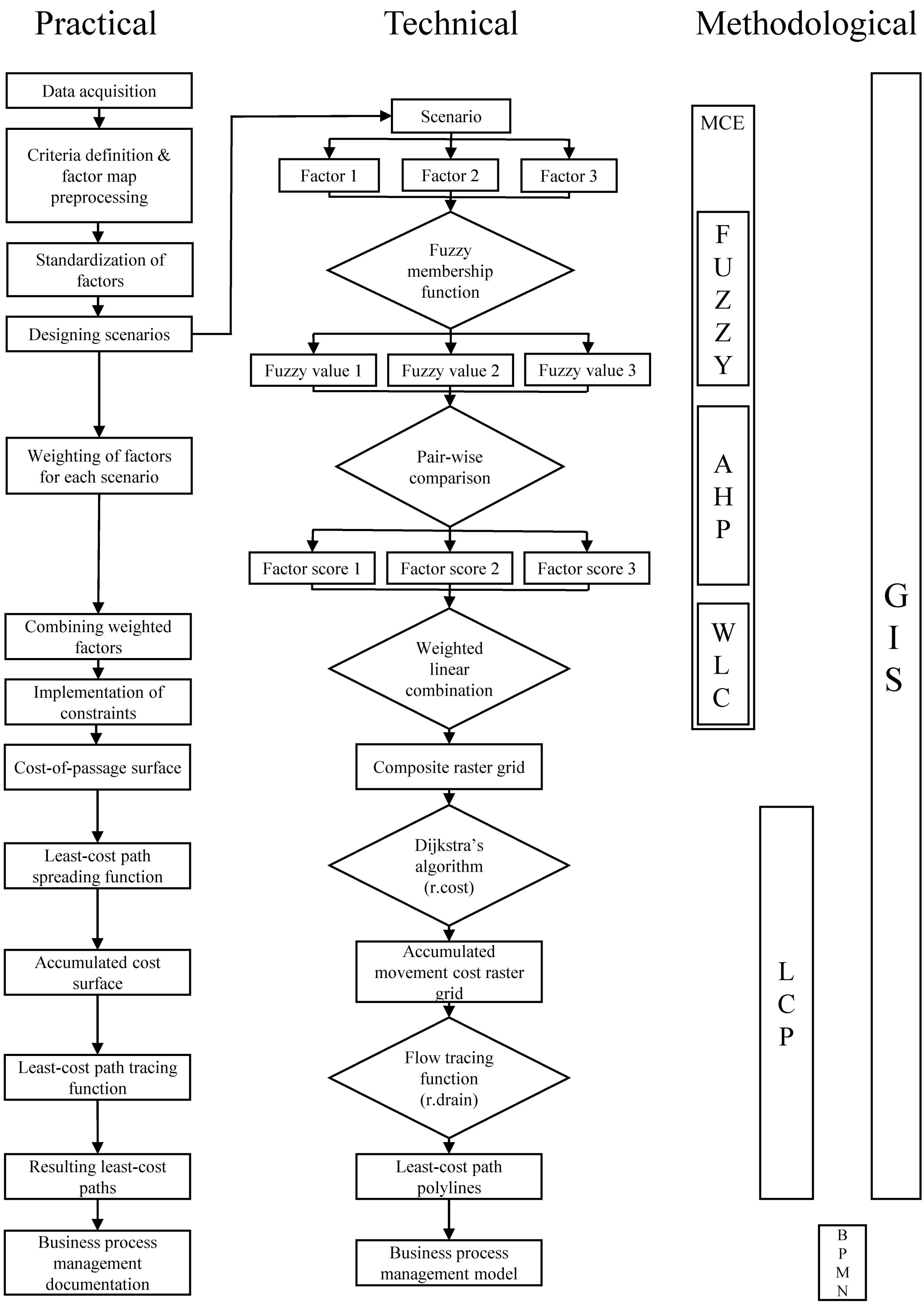


Figure 14. Flow diagram of overall research workflow.

4.1 Datasets

This study was based on the use of two types of spatial data: planning data and physiographic data (Table 1). As planning data is based on restrictions set on the study area, physiographic data forms the physiographic basis for least-cost path analysis through defining the frictions and impedance to movement. Therefore, the planning data can be seen as a framework for the analysis. While physiographic data offers accurate, remotely sensed data of the environment, planning data is based on human assumptions and evaluations in the form of predefined areas and locations that are determined suitable for construction operations.

4.1.1 Planning data

Planning data consists of information presented in the master plan of the project (Kemiönsaari 2014) or material which is one way or another relevant for planning. These data form the overall operative framework for the study. Project area borders represent the area that is designated as the planning area in the official master plan. This area is dedicated to the construction and production of wind power and its infrastructure. The official project area borders presented in the master plan were slightly adjusted for the analysis phase of this study. The existing road near turbines N01, N02 and N03 has been included in the analysis, as it will very likely be utilized in the project. It is not specifically forbidden to build infrastructure outside the project area, in this case it has to be agreed between the land owner and the project developer company. Also, four different entrance points have been defined near the project area borders to provide the main roads leading into the project area.

All the wind power plant locations were already defined in the study area. Fixed coordinates for each wind power plant and the locations of wind power areas in which the construction area will be planned, were obtained from the master plan. The construction areas inside wind power areas are not specifically defined in the master plan, but in this study, a preliminary construction areas of a generic size and shape was used.

In order to plan collector system alignment, the potential locations for electrical substations or PCCs have to be defined in order to know the starting points for pathway optimization. Preliminary substation position is also determined in the master plan. In addition, preliminary positions for cable cabinets are provided. These cabinets act as hubs for cabling and collecting the cables from groups of up to five turbines. Every cabinet is then separately connected to the main substation.

Additional planning data that is also defined in the master plan includes living habitats for protected species, such as capercaillie or flying squirrel (Kemiönsaari 2014). In these areas, the planning of wind power areas and infrastructure is restricted. This is usually only an issue during micro-siting, because the protected species living habitats are taken into account already during the earlier planning phases such as the EIA process (EFE 2012). Restriction areas dataset also includes agricultural fields and water bodies. Fields are not specifically described in the master plan as forbidden for construction of infrastructure, but avoiding them has been mutually agreed between EFE and the land owners. Some of the water bodies have been declared in the master plan as ecologically protected ponds and thus restricted for construction. All other water bodies in the project area were merged with this data, as constructing expensive crossings of water bodies in the infrastructure alignment planning would be unnecessary.

Furthermore, in order to evaluate the results of infrastructure alignment planning, a reference is required. Data about preliminary planned roads described in the master plan (Kemiönsaari 2014) can be used to compare the results of the road path optimization process. The preliminary plan for the internal grid is used only as a reference in a numerical length comparison with the produced grid optimization results.

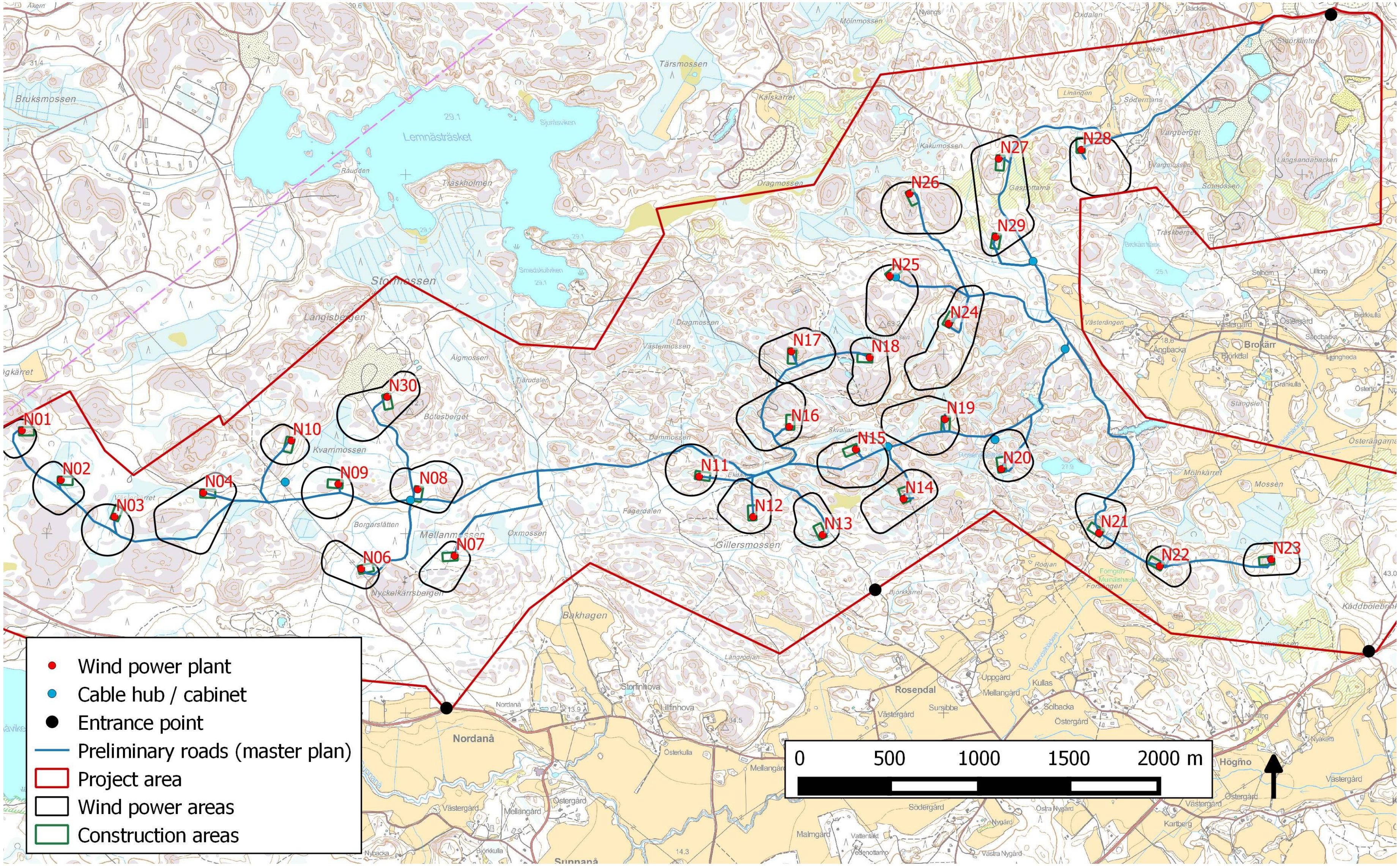


Figure 15. Planning data and preliminary roads presented in the master plan (Kemiönsaari 2014).

4.1.2 *Physiographic data*

Most of the physiographic data used in this study is based on a lidar dataset. This data includes Digital Terrain Model (DTM), Digital Surface Model (DSM), water bodies, slope map and hillshade map. These datasets were derived from an airborne lidar scanning operation ordered by EFE Ab. The scanning was conducted in April 2015 and the result data was delivered in point cloud -based LAS format which uses a ground type classification. The classes that were utilized in this study were elevation data of ground level, vegetation and areas recognized as water bodies. The scanning was conducted with a mean point density of 7–11 points per square meter. The estimated horizontal accuracy of the data is about 0.14 meters and the estimated vertical accuracy is 0.08 meters.

In Finland, NLS is currently conducting airborne lidar scanning operations across the whole country and it is estimated to be completed in 2019. The nationwide lidar data (Paikkatietohakemisto 2016) will have a mean point density of 0.5 points per square meter and the elevation model will have horizontal accuracy of 2.0 meters and vertical accuracy of 0.3 meters. The data is offered for use free of charge. However, at the time of this study, lidar data by NLS was not yet available for the study area of this thesis. It could have been possible to use the old elevation model by NLS, but because of the horizontal accuracy of 10 meters and vertical accuracy of 1.4 meters, this data was considered insufficient regarding resolution to be used for this study. The old NLS elevation model (Paikkatietohakemisto 2015) has been produced by combining various different data such as verified elevation contours, digitized ground points and terrain database objects with an elevation attribute. Due to the utilization of lidar devices, the new NLS elevation model does not only offer superior resolution to the old NLS elevation model, but it also outperforms it in data consistency.

In order to successfully integrate the wind park access road network into the national road network, data describing existing roads is required. *Digiroad* is an extensive dataset produced by the Finnish Transport Agency (FTA 2016). It contains information about Finnish roads and their locations. In this study, the 2012 version of Digiroad is used and treated as a physiographic dataset, because the features of existing roads are approached from a physical perspective, emphasizing the leveled, even and solid ground of roads. The usage of existing roads in this study is based on the argument that new roads are significantly easier to construct on already existing roads. Even if not up to standards set by wind turbine transportation operations, even a narrow forest road can already offer a huge boost for road construction.

Table 1. Datasets of the study consist of planning data and physiographic lidar data.

Dataset	Description	Optimization goal	Date	Resolution	Type	Source
DTM	Digital Terrain Model describing ground level elevation	Take the directional component of slopes into account	2015	Horizontal: 1,00 m Vertical: 0,18 m	Raster	Lidar data (EFE Ab)
DSM	Digital Surface Model describing elevation of vegetation	Minimize crossing areas with high vegetation	2015	Horizontal: 1,00 m Vertical: 0,18 m	Raster	Lidar data (EFE Ab)
Water bodies	Water bodies derived from DSM	Minimize crossing water bodies	2015	Horizontal: 1,00 m Vertical: 0,18 m	Raster	Lidar data (EFE Ab)
Slope map	Slope gradients in degrees, derived from DTM	Minimize high slope gradients on route	2015	Horizontal: 1,00 m Vertical: 0,18 m	Raster	Lidar data (EFE Ab)
Hillshade map	Shaded relief map derived from DTM	Visualize results	2015	Horizontal: 0,10 m Vertical: 0,18 m	Raster	Lidar data (EFE Ab)
Wind power master plan	Overall land use inside the project area	Set a framework for planning	2014	1:10 000	Image	Kemiönsaari municipality
Project area borders	Borders of the planning area	Stay inside the planning area	2014	1:1 000	Vector	Master plan
Construction areas	Construction areas for power plants (generic design)	Maximize the utilization of leveled construction areas	2015	1:1 000	Vector	Turbine manufacturers
Restriction areas	Restriction areas for planning, including fields and protection areas	Minimize proximity to protection areas / Exclusion areas (constraint)	2014	1:1 000	Vector	EIA (EFE Ab)
Entrance points	Locations for pre-planned transportation entrance points	Starting points for LCP analysis	2015	1:1 000	Vector	EFE Ab preliminary planning
Substation & cable hub locations	Locations for planned substations and cable hubs/cabinets	Starting points for LCP analysis	2014	1:1 000	Vector	Master plan & EFE Ab (hubs)
Power plant locations	Locations for planned wind power plants	Stopping points for LCP analysis	2014	1:1 000	Vector	Master plan
Digiroad (existing roads)	Finnish road network	Maximize utilization of existing road network	2012	1:1 000	Vector	Finnish Transport Agency
Preliminary roads	Preliminary planned roads for a wind power project	Compare different layouts	2014	1:1 000	Vector	Master plan

4.1.3 Data preprocessing

All relevant LAS data from lidar scanning was converted into a raster-based regular grid models with a 1 meter pixel size. The data would have allowed a maximum resolution of 0.15 meters, but when running initial tests, the processing times proved out to be extremely long even for a high-end consumer-level PC. 1 meter pixel size was a compromise between processing times and relevant accuracy for the purposes of this study. Even with the resolution of 1 meter, the amount of pixels was very high due to the large area coverage. Therefore the project area was split in four parts, then converted and then merged into a one large grid.

The conversion was carried out using triangulation-based interpolation type with the nearest neighbor method. With this method, the cells will be assigned with values obtained by interpolating measurements from a triangulated representation of the LAS dataset (ArcGIS Resource Center 2012). Triangulation always fills the void areas on a grid with true interpolation and the natural neighbor method always uses the natural neighbor value to determine cell value (ArcGIS Resource Center 2012). The triangulation process was carried out only for the DTM, as a consistent and a whole terrain elevation raster grid was required for the analysis. For DSMs describing vegetation and water bodies, binning-type conversion with nearest neighbor method was used in order to store only actual data occurrences (ArcGIS Resource Center 2012). In binning method, it is possible to leave areas that have a data void, empty. Because water and vegetation are not necessarily recorded along the whole grid surface, there may be areas with zero occurrences of these classes. Thus, interpolation by triangulation would distort the data in areas with no data.

After the conversion operations, a slope analysis was conducted using the DTM as the input data. In the resulted slope raster, each pixel in the raster surface was assigned with the slope value of the direction of the steepest descent. Thus it reflects the changes in the topography along the surface. Additionally, a hillshade map was also produced from the DTM dataset to help in the visualization of results. The raster datasets describing vegetation and water bodies contain information about the total height above sea level. To derive relative values for these datasets, the DTM raster was subtracted from vegetation and water body raster datasets.

Because LCP analysis uses raster grids for modeling, multiple data type conversion operations were required. Natural protection areas, overall restriction areas, existing roads and construction areas were originally in a vector format, but for the purposes of this study, these datasets were all converted to raster format. In addition, to be able to use the data in GRASS GIS software, all the raster files had to be imported separately from QGIS view to the file format used by GRASS GIS. In every analysis of every scenario, a pixel size of 1 meter was used.

4.2 Multi-criteria evaluation methodology (MCE)

In this study, multi-criteria evaluation is used to derive realistic cost-of-passage surfaces for LCP analysis by logically combining the established methods of fuzzy logic, AHP and WLC. In practice, the MCE methodology presented in this study begins with standardization and then multiplying the standardized criteria with the outcome weights from the pair-wise comparison. The process continues by summing up each alternative to create a ranking-based composite raster map. Final step is to multiply the resulting composite map with a Boolean constraint map if necessary. In this study, the steps for producing a cost-of-passage surface for LCP analysis are:

1. Criteria definition and factor map preprocessing
2. Factor standardization in accordance to reversed fuzzy logic presented by Jiang & Eastman (2000)
3. Designing relevant scenarios according to desired optimization objectives
4. Pair-wise comparison using Analytical Hierarchy Process (AHP) developed by Saaty (1977) to derive the weights
5. Combining factor maps for each scenario using Weighted Linear Combination (WLC) practices discussed by Malczewski (2000)
6. Boolean overlay multiplication to implement constraints

This 6-step process is similar to those used by Atkinson et al. (2005) and Choi et al. (2009). Atkinson et al. used a very similar methodology in their study that combined MCE and LCP for road planning, but with the exception of using a fuzzy gamma operator in combining factors. Choi et al. used a nearly identical methodology as presented in this study, but were able to implement accurate mathematical calculations derived from literature in the factor standardization phase to produce fuzzy values less prone to subjective assumptions.

4.2.1 Criteria definition & factor standardization

The selected criteria should reflect the objectives that the planner desires to achieve. The path optimization problem presented in this study utilizes least-cost path analysis to find the optimal paths in order to minimize resources used in infrastructure construction. Properties of attributes that are relevant to road and electric grid construction have been selected and used for evaluating criteria selection options. In this study, five different criteria were used. Of these five criteria, four were factors and one was a constraint (Table 2).

The criteria selected for this study were standardized using fuzzy membership functions customized for pathway construction and LCP analysis purposes. The objective was to study the degree of membership in the set defined as *suitable* for road construction. In practice, the raw data values of a raster grid were reclassified according to reverse fuzzy logic in order to create a new raster grid in which the cell values represent suitability for road construction (Table 2). The fuzzy values presented were determined by following the same fuzzy logic described in the MCE-oriented literature (Eastman 1999; Jiang & Eastman 2000) and fuzzy set theory literature (Bonham-Carter 1994; Zimmermann 2010). Also, as the assumptions inspired by engineering evaluations and other LCP-related literature, such as Yu et al. (2003), Atkinson et al. (2005) and Choi et al. (2009), were transformed into discrete quantitative values, subjective evaluation was inevitably required. In this case, engineering evaluations refers to multiple classified info sheets for road requirements provided by turbine manufacturers. According to these evaluations, 5–7 degrees is the general threshold for maximum slope gradient in the transportation of a large, modern wind turbine.

Table 2. Fuzzy value classification for each factor (slope gradient, vegetation height, protection area buffer, distance from roads & construction areas) with raw data value ranges and re-classified fuzzy values.

Class	Slope gradient (degrees)	Fuzzy value	Class	Vegetation height (m)	Fuzzy value
1	0–8	0.1	**1**	0	0.1
2	8–12	0.3	**2**	0–1	0.2
3	12–16	0.4	**3**	1–5	0.3
4	16–30	0.6	**4**	5–10	0.4
5	30–45	0.8	**5**	10–15	0.5
6	45–90	1.0	**6**	15–20	0.6
-	-	-	**7**	20–30	0.8

Class	Protection area buffer (m)	Fuzzy value	Class	Distance from roads & construction areas (m)	Fuzzy value
1	> 60	0.1	**1**	0	0.1
2	60–40	0.2	**2**	> 0	0.5
3	40–20	0.4	-	-	-
4	20–0	0.5	-	-	-
5	0	1.0	-	-	-

Factor: slope gradient in degrees

Slope gradient determines the level of groundworks, such as excavation and demolition works needed, in order to construct a road that is suitable for transportation. Groundworks are costly and have a direct effect on local environment. The values that were used to estimate the importance of slope gradients for road construction are based on evaluations by turbine manufacturers and practices used in similar LCP-related studies such as Yu et al. (2003) and Atkinson et al. (2005). One vital aspect in the estimations was the generally acceptable maximum road gradient of 5–7 degrees, which was used as a rough guideline. In generating fuzzy values, a range of 0–8 degrees was selected to determine the most suitable (0.1) slope gradients for road construction. The range was kept slightly broader in an attempt to prevent too strict classification. As areas with slope gradient of under 8 degrees are defined as most suitable for wind turbine transportation, there is a small jump in the fuzzy values between classes 1 and 2 to emphasize the particular suitability of class 1. It was assumed, that an area with a slope gradient of over 8 degrees definitely requires at least moderate amount of groundworks. Gradient of 45 degrees was selected as the high anchor point for the fuzzy set, as a slope already this steep requires very much resources to transform into a road that corresponds to wind turbine transportation requirements. As slope gradient increases past this point, it is difficult to assume that the amount of required resources would continue to grow incrementally.

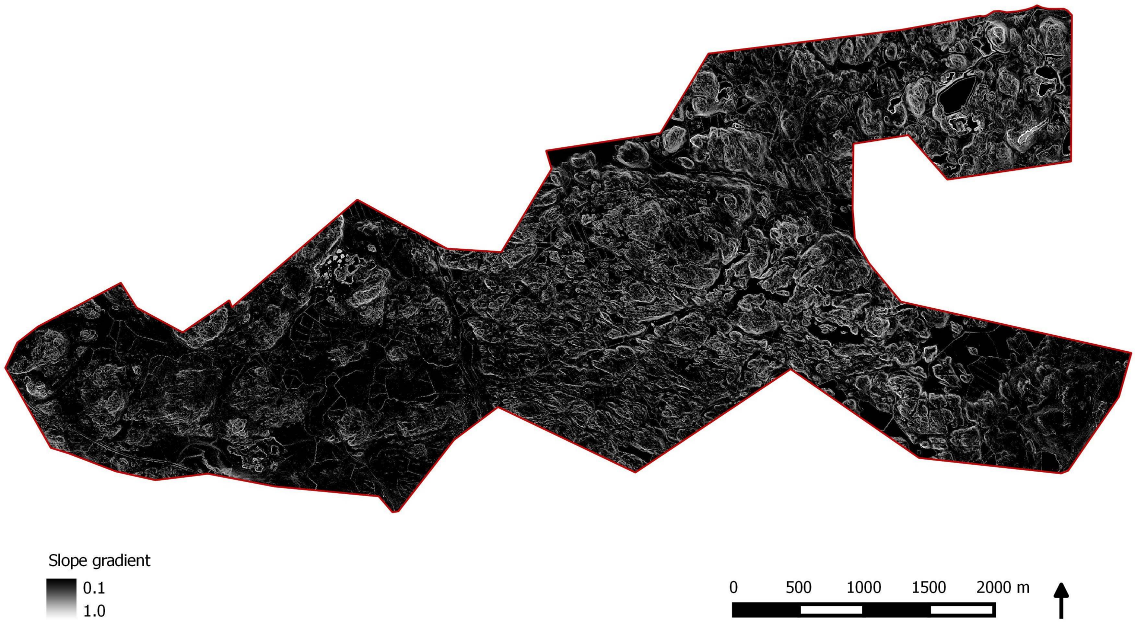


Figure 16. Slope gradient in degrees presented as a standardized factor map. Lower value (darker tone) indicates more suitable areas for road construction.

Factor: vegetation height in meters

The amount of vegetation determines the amount of clearing works required in order to construct a road. The costs are generally calculated in cubic meters of cleared material. Sudden clearing of vegetation is not the most important factor cost-wise, but it can have a local effect on environment and ecosystems. Ecologically, taller vegetation generally indicates older vegetation. In addition to economic aspects, an assumption was made to regard areas with taller vegetation also ecologically as more valuable.

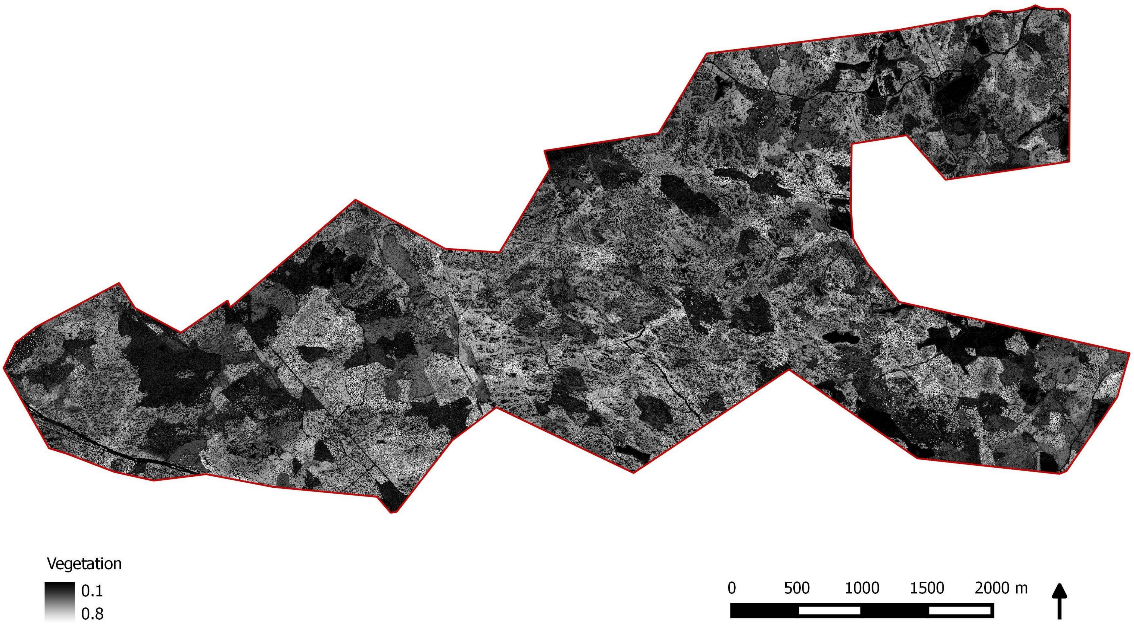


Figure 17. Vegetation height in meters presented as a standardized factor map. Lower value (darker tone) indicates more suitable areas for road construction.

Factor: distance from natural protection areas

Natural protection areas are defined in order to protect valuable natural areas, which might contain habitats for endangered species or other natural phenomena that is considered special. These areas are usually a result of an extensive environmental impacts assessment program and thus marked in the master plan as restricted areas for construction. In this study, buffer zones up to 60 meters were created around these areas in order to gradually increase the movement costs in the close proximity of these areas.

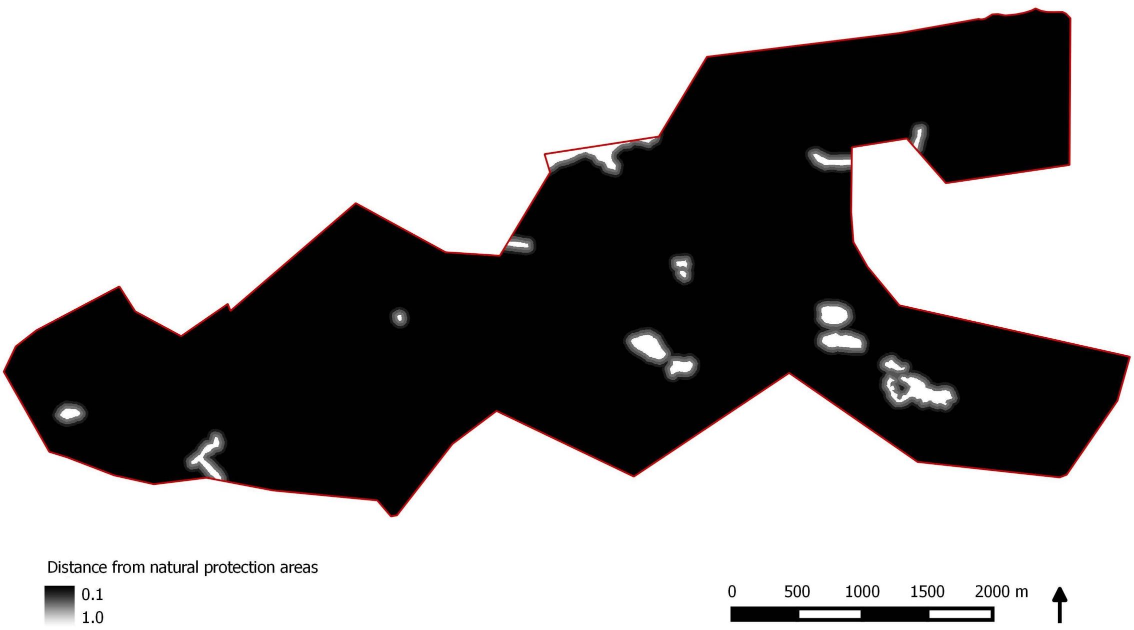


Figure 18. Distance from natural protection areas presented as a standardized factor map. Lower value (darker tone) indicates more suitable areas for road construction.

Factor: distance from construction areas and existing roads

As discussed in Chapter 2, upgrading existing roads can be significantly less expensive than building a completely new road. Thus, existing road network can be seen as an important criterion in the path optimization process. Given the predefined construction areas, a generic design of a construction platform was used to indicate the areas, which will be leveled in any case during the power plant construction phase. These areas were logically combined with the existing roads to represent areas with lower costs for movement. This factor is slightly overlapping with slope gradient factor, because existing roads tend to naturally have a smaller slope gradient.

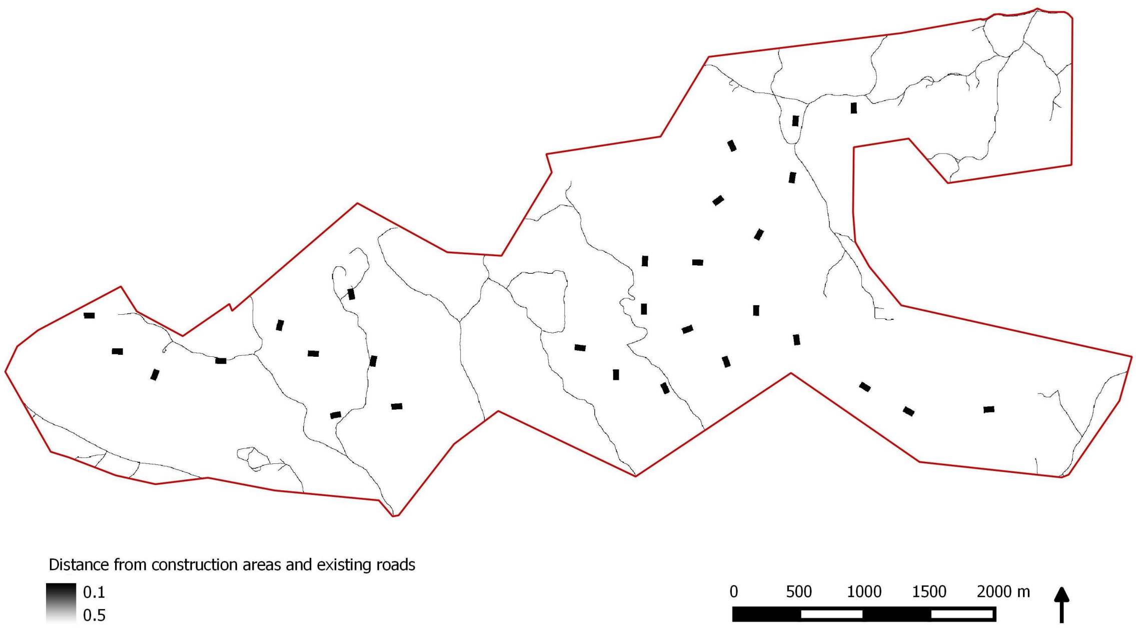


Figure 19. Distance from construction areas and existing roads presented as a standardized factor map. Lower value (darker tone) indicates more suitable areas for road construction.

Constraint: restriction areas defined in the master plan

All the contents of the factor describing distance from natural protection areas were also in the restriction areas constraint as well. Additionally, restriction areas included all missing water bodies, archeological findings and mating grounds of capercaillie. This constraint was used as a Boolean overlay map, with the no-go areas valued at 0 and other areas at 1. After the Boolean multiplication operation, all the cells values at 0 were transformed into null-values with the *r.null*-module in GRASS GIS in order to exclude them from the LCP analysis.

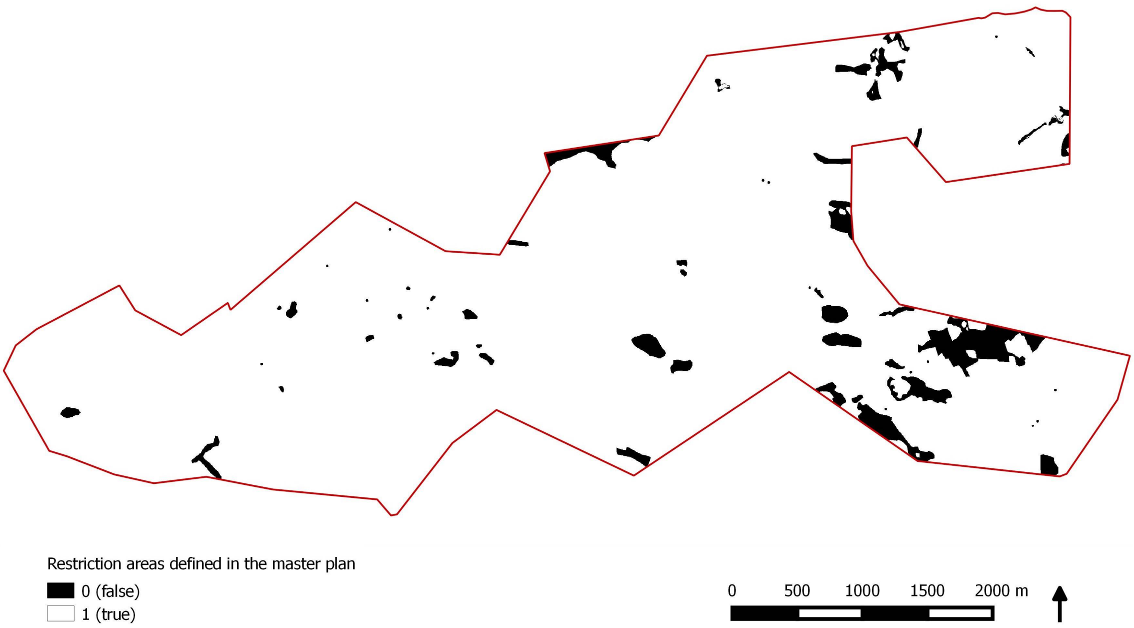


Figure 20. Restriction areas defined in the master plan presented as a Boolean constraint map. Value zero (false) indicates restricted areas for road construction.

4.2.2 Criteria weights & optimization scenarios

In this study, the standardized fuzzy values describing suitability were kept constant through the whole optimization process. Instead, the weights for different criteria were separately re-evaluated for the optimization process of each scenario by using pair-wise comparison of AHP. The weighting process was carried out in QGIS by using EasyAHP-plugin. First, the relative importance of each factor was determined in a pair-wise comparison matrix according to the 9-point reciprocal scale by Saaty (1977) to derive the weights of each factor in a scenario. The pair-wise comparison matrices are presented for scenarios R2, R3, and G1/G2 in Figures 22, 24 and 26 respectively. AHP indicators are presented as well, with lambda-value (λ) indicating the eigenvalue of the matrix and *Consistency Index* (CI) representing the consistency of judgements across all factor comparisons. The Consistency Ratio (CR) of pair-wise comparisons is calculated by comparing CI and *Random Consistency Index* (RI). In this study, CR remained under the acceptable limit of 0.100 in every scenario.

After deriving the weights, each criterion was multiplied with its respective weight and all the criteria were summed up to form a factor map for a single scenario. Finally, if required by a scenario, a constraint was added to the factor map by performing a Boolean multiplication on the factor map to remove restriction areas from consideration. The factor maps and constraints were combined with *r.mapcalc*-module in GRASS GIS. The pair-wise comparison tables and resulted factor maps are presented separately under each scenario.

Different scenarios with different objectives and criteria have been planned and produced in order to evaluate results of LCP analysis for varying planning objectives (Table 3). There are five scenarios in total: three scenarios for roads and two scenarios for electric grid.

Table 3. Optimization scenarios and objectives and their respective criteria and constraints.

Scenario	Target	Optimization objective	Criterion 1	Criterion 2	Criterion 3	Constraint
R1a	Roads	Minimize demolition works (Isotropic spread)	Slope (absolute)	-	-	-
R1b	Roads	Minimize demolition works (Anisotropic spread)	Slope (absolute)	-	-	-
R2	Roads	Minimize impact on nature (Isotropic spread)	Slope	Vegetation	Protection areas	-
R3	Roads	Minimize road costs (Isotropic spread)	Slope	Construction areas & existing roads	Vegetation	Planning restrictions
G1	Grid	Minimize grid costs (Isotropic spread)	Slope	Construction areas & existing roads	Vegetation	Planning restrictions
G2	Grid	Minimize grid costs when road planning is prioritized (Isotropic spread)	Slope	Construction areas & existing roads & R3 results	Vegetation	Planning restrictions

Scenario R1

First road scenario R1 was the only scenario with only one criteria, slope gradient. The objective was to minimize demolition- and excavation works. Thus the optimization results could be described as the geotechnical optimum. Scenario R1 was divided into two parts: R1a used algorithm with isotropic spread and R1b used anisotropic spread. The aim of dividing scenario R1 into two parts was to study how large of an effect does the directional component of anisotropic spread have on the path optimization results. Additionally, because scenario R1 only contains one criterion, the slope gradient map was directly used as the cost-of-passage map. This means that no standardization procedures were required to be carried out for this scenario and thus the absolute slope gradient values were used straightforwardly as impedance to movement.

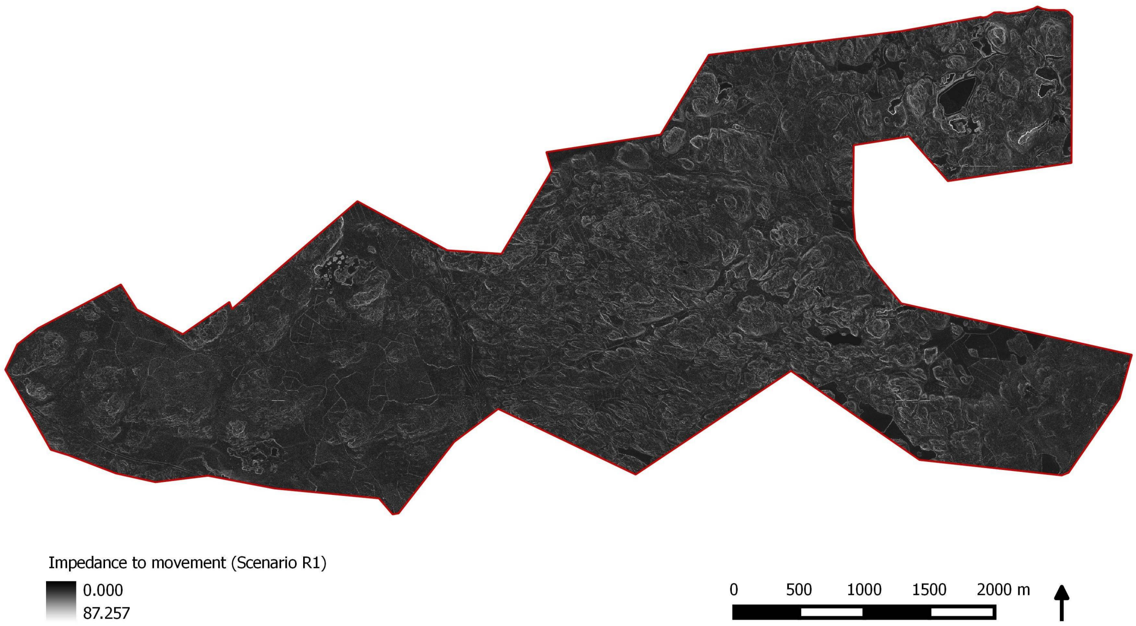


Figure 21. Cost-of-passage composite map for scenario R1.

Scenario R2

In R2, slope gradient, vegetation height and distance from natural protection areas were combined in order to create a scenario to minimize the impact on nature. Natural protection areas were heavily weighted in order to emphasize their more sensitive status from ecological perspective.

	Slope	Natural Protection	Vegetation
Slope	1	0.5	2.0
Natural Protection	2.0	1	4.0
Vegetation	0.5	0.25	1

AHP Indicators

λ = 3.0

CI = 0.0

CR = 0.0

Layer Weights

	Layer Name	Weight
1	Slope	0.286
2	Natural Protection	0.571
3	Vegetation	0.143

Figure 22. Top: pair-wise comparison matrix according to Saaty's 9-point reciprocal scale and corresponding AHP indicators in scenario R2. Bottom: derived weights for each factor.

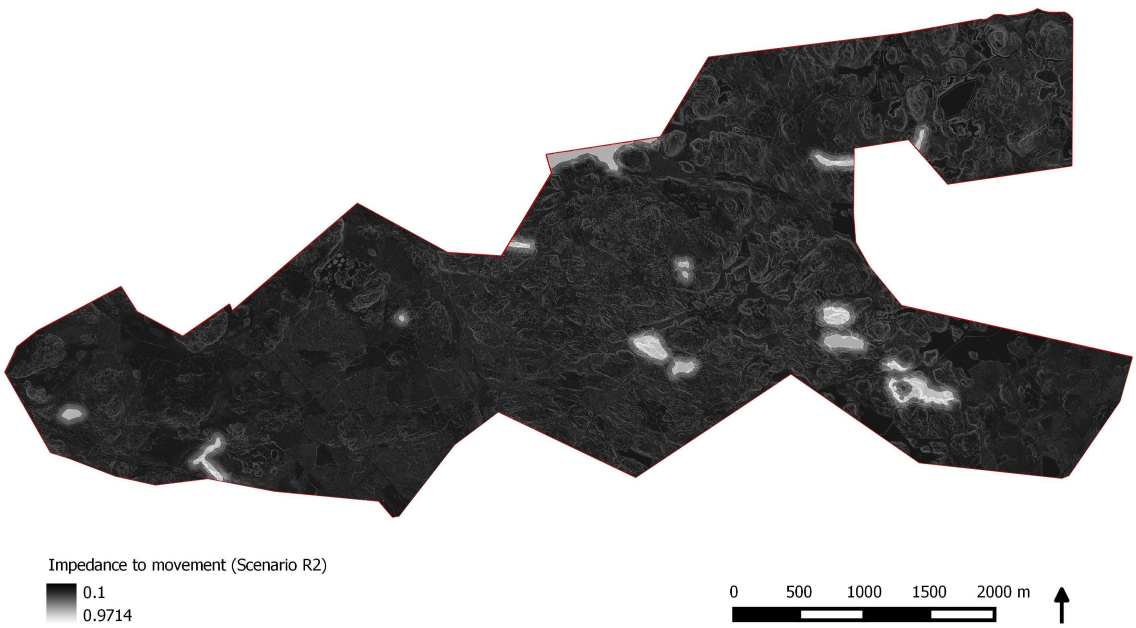


Figure 23. Cost-of-passage composite map for scenario R2.

Scenario R3

Scenario R3 aimed to optimize road construction costs with the combination of slope gradient, distance from construction areas and existing roads and vegetation height. Additionally, realistic planning restrictions were implemented in the scenario as a Boolean constraint. Overall, scenario R3 emphasized construction areas and existing roads in a relatively heavy manner. Three assumptions were made in terms of cost structure:

- Utilizing existing roads halves the road construction price
- Demolition- and excavation works are very much more expensive than costs from clearing of vegetation
- Utilizing existing roads is equally important as minimizing slope gradient

	Slope	CAreas & Roads	Vegetation
Slope	1	1.0	7.0
CAreas & Roads	1.0	1	7.0
Vegetation	0.143	0.143	1

AHP Indicators

λ = 3.003

CI = 0.002

CR = 0.003

Layer Weights

	Layer Name	Weight
1	Slope	0.467
2	CAreas & Roads	0.467
3	Vegetation	0.067

Figure 24. Top: pair-wise comparison matrix according to Saaty's 9-point reciprocal scale and corresponding AHP indicators in scenario R3. Bottom: derived weights for each factor.

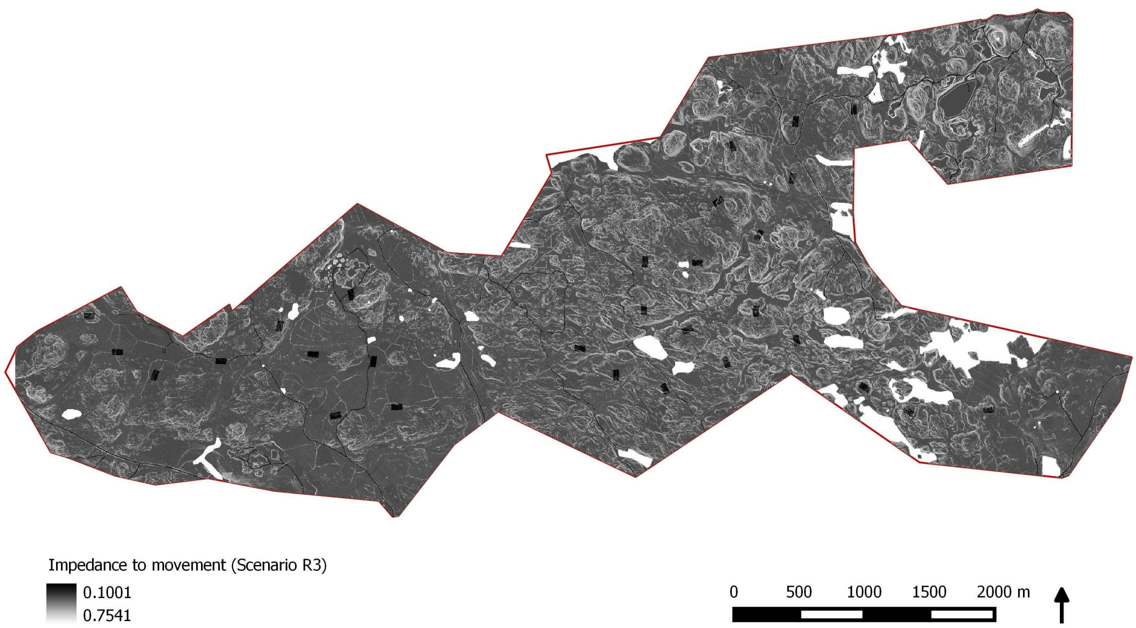


Figure 25. Cost-of-passage composite map for scenario R3.

Scenario G1

Scenario G1 was the cost optimization scenario for the alignment of electric grid. It used the same criteria and constraint as R3. The general practice in the industry has been to first construct the roads and then align the cables in ditches alongside the roads. Road planning has been generally prioritized to minimize the effects on landscape. However, depending on the construction method, underground electric grid usually has significantly smaller effect on the overall landscape. Cable ditches are also narrower and thus easier to construct than roads. In some cases, cables can even be installed on the ground, covered with steel pipes. In scenario G1, construction areas and existing roads have notably lighter weights than in scenario R3. The assumptions for the costs are following:

- Aligning cables alongside roads is not mandatory: lighter weights on construction areas and existing roads
- Demolition- and excavation works are very much more expensive than costs from clearing of vegetation
- Minimizing slope gradient is very much more important than utilizing existing roads

	Slope	CAreas & Roads	Vegetation
Slope	1	7.0	7.0
CAreas & Roads	0.143	1	2.0
Vegetation	0.143	0.5	1

AHP Indicators

λ = 3.055

CI = 0.028

CR = 0.048

Layer Weights

	Layer Name	Weight
1	Slope	0.767
2	CAreas & Roads	0.143
3	Vegetation	0.09

Figure 26. Top: pair-wise comparison matrix according to Saaty's 9-point reciprocal scale and corresponding AHP indicators in scenarios G1/G2. Bottom: derived weights for each factor.

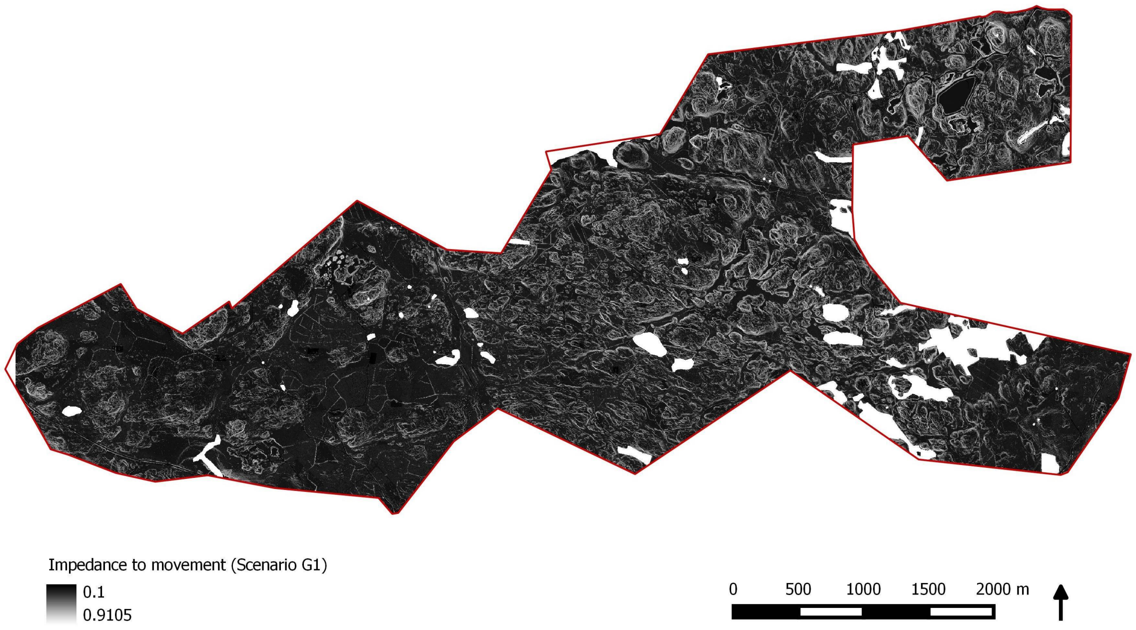


Figure 27. Cost-of-passage composite map for scenario G1.

Scenario G2

In scenario G2, the cost-optimized roads from scenario R3 were integrated into the factor map of construction areas and existing roads. This scenario aimed to simulate the generally established workflow of prioritizing road planning over grid planning. Additionally, it was expected for the newly implemented roads to moderately control the simulated paths for cables. The assumptions for costs correspond to the assumptions made for scenario G1. Thus, the weights were also similar to the ones presented for G1.

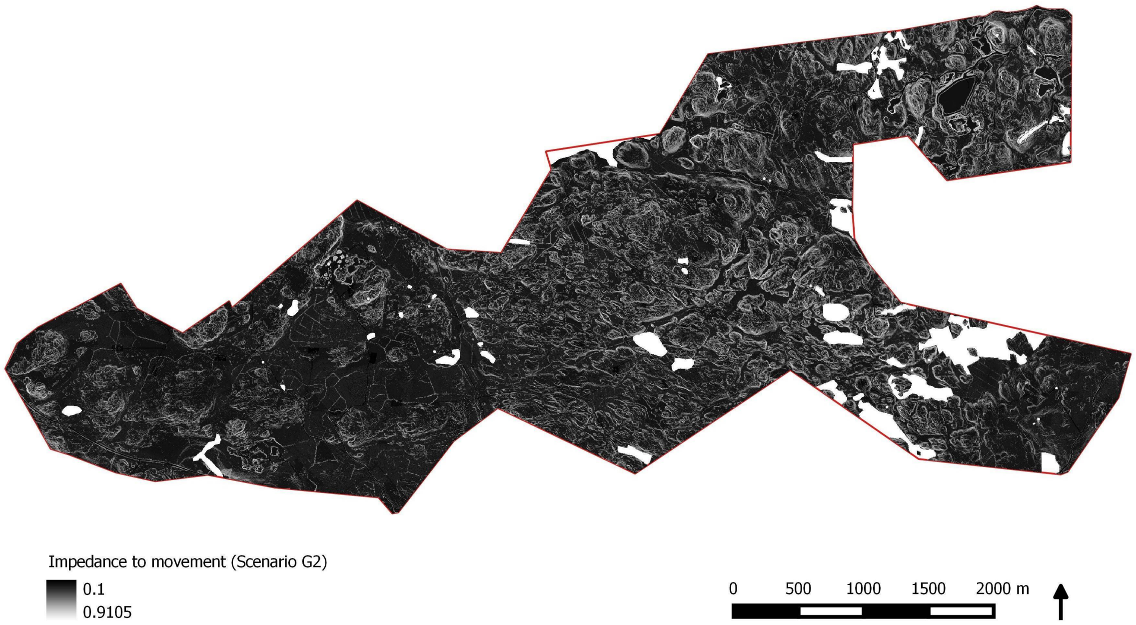


Figure 28. Cost-of-passage composite map for scenario G2.

4.3 Least-cost path analysis (LCP)

4.3.1 Analysis parameters

The least-cost path analysis begins with the creation of an accumulated cost surface from a cost-of-passage surface (Douglas 1994). The second step is to trace down the path which has the lowest accumulated cost between starting point and ending point along the accumulated cost surface (Douglas 1994). In this study, when modeling the access road network, the study area was logically divided into four clusters: western (W), central (C), eastern (E) and northern (N). The clusters were determined according to four predetermined entrance points (Figure 29) to the project area. Every scenario was modeled in four parts, as each cluster was modeled separately. In each cluster, the entrance point coordinates were entered as the starting point for the algorithm, and the coordinates of all the turbines in a cluster were entered as the stopping points for the algorithm. For scenario R3, two separate connection paths were modeled afterwards in order to connect the clusters to each other to form a unified access road network.

In electrical grid modeling, six predetermined cable hub locations (Figure 25) were used as the starting points and a group of up to five turbines forming six separate *circuits*, as the coordinates for the stopping points. Additionally, every cable hub was separately connected to the substation.

The LCP modeling procedures were carried out in GRASS GIS, which has three separate integrated modules to conduct a least-cost path analysis. Module *r.cost* straightforwardly uses a cost-of-passage surface to create an accumulated cost surface with isotropic spread. In addition to using cost-of-passage surface, the second module, *r.walk*, requires the elevation data for analysis in order to calculate the directional component for slopes (GRASS GIS manual 2015). Thus, *r.walk* produces an accumulated cost surface with anisotropic spread. A decision was made to use the isotropic *r.cost* -module as the primary tool in this study. This decision was based on literature review, experiences from scenario R1, and the more flexible overall nature of *r.cost* compared to *r.walk*. After extensive preliminary simulations prior to carrying out the actual study, it could be concluded that isotropic and anisotropic spread functions produce quite similar results in this study area.

Both modules, *r.cost* and *r.walk*, use Dijkstra's algorithm for computing the minimum cumulative costs and additionally, both modules can be run with *Knight's move* –flag (GRASS GIS manual 2015). When entered in the command parameters, the algorithm uses the Knight's pattern -method in identifying neighboring cells. Due to its advantages presented by Yu et al. (2003), every modeling procedure in this study was conducted with the Knight's pattern.

The third module which was used in this study is *r.drain*. It has been developed to trace down a flow through a least-cost path in an elevation model. However, it may very well be used to trace down a cumulative cost map generated with *r.cost* or *r.walk*. *R.drain* outputs an integer cell raster map with value 1 along the least-cost path and value null elsewhere (GRASS GIS manual 2012). Final steps of the analysis were to convert drain raster map into vector lines and then join each of these cluster lines into one vector dataset.

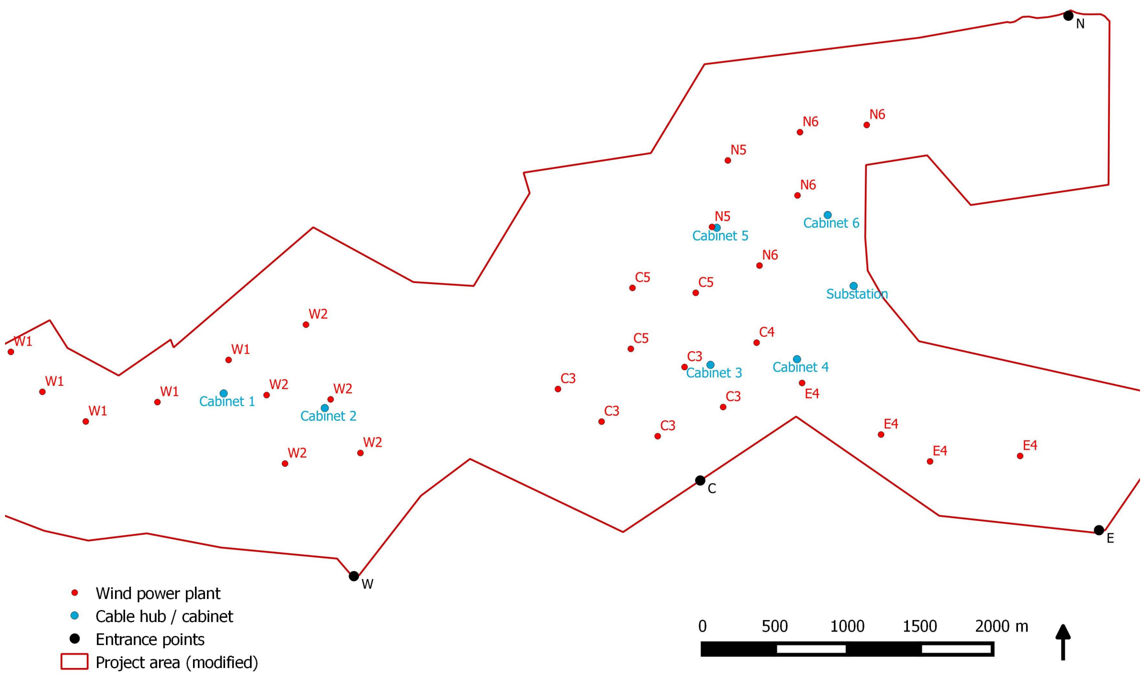


Figure 29. Entrance points for road scenarios and cabinets & substation for grid scenarios. Character indicates road scenario starting point, number indicates grid scenario starting point.

4.3.2 Accumulated cost surfaces

Scenario R1

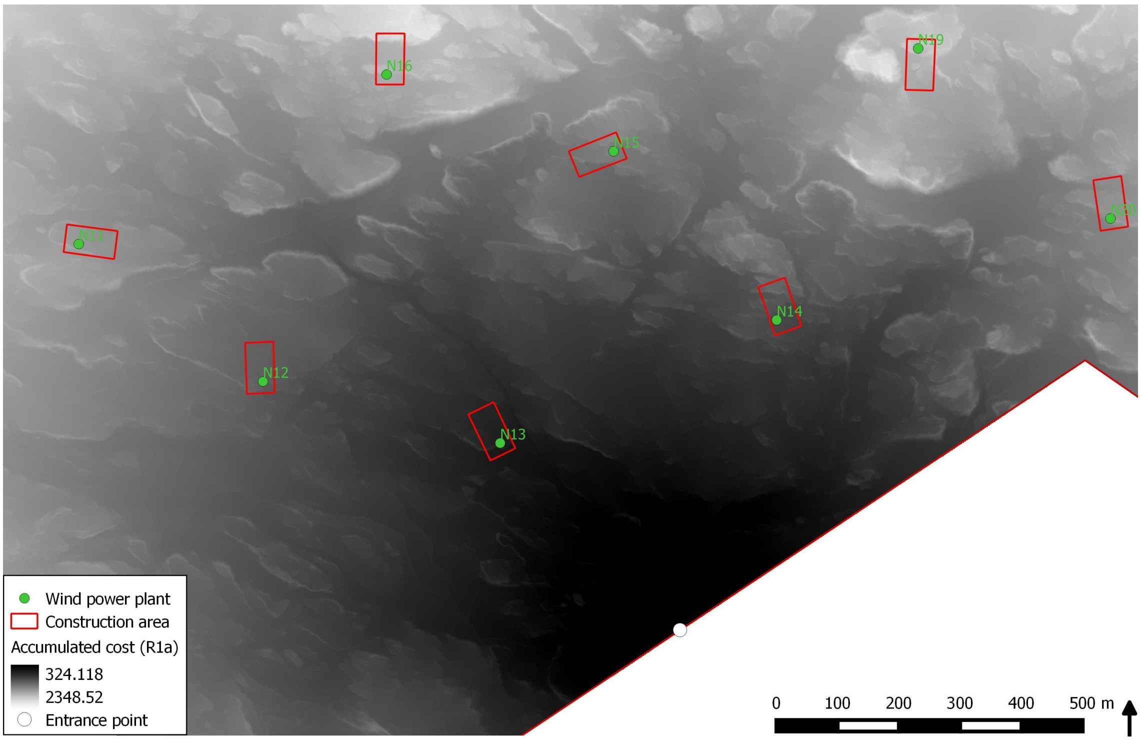


Figure 30. Scenario R1a accumulated cost surface in the central cluster.

Both R1 scenarios, R1a and R1b, use only the slope map as the cost surface. Thus, it is logical that the resulted accumulated cost surface has similar slope patterns than the input slope data. Only R1a is presented here, because there was no visually distinguishable differences between the accumulated cost surfaces of R1a and R1b.

Scenario R2

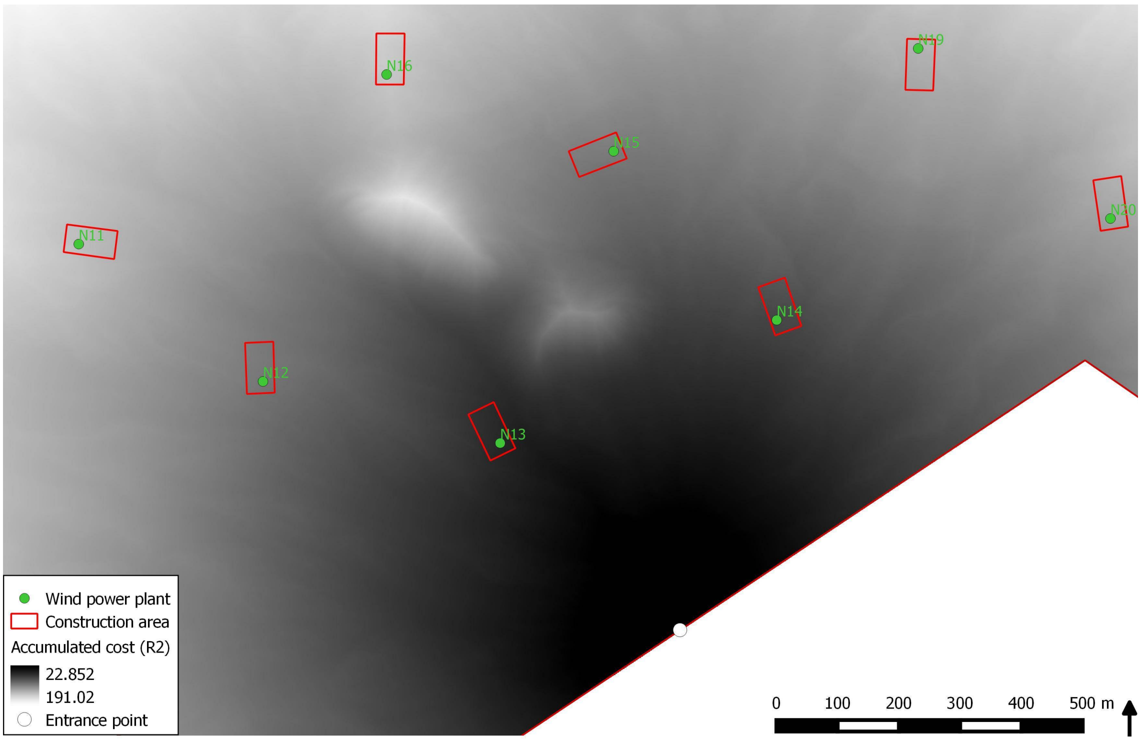


Figure 31. Scenario R2 accumulated cost surface in the central cluster.

In the accumulated cost surface of R2, the heavily weighted natural protection areas are clearly visible. The use of buffer zones is also evident, as the borders of these protection areas are smooth and almost linearly fading. Otherwise, the spreading function is quite balanced in every direction, following low-slope areas, especially the road heading north-west.

Scenario R3

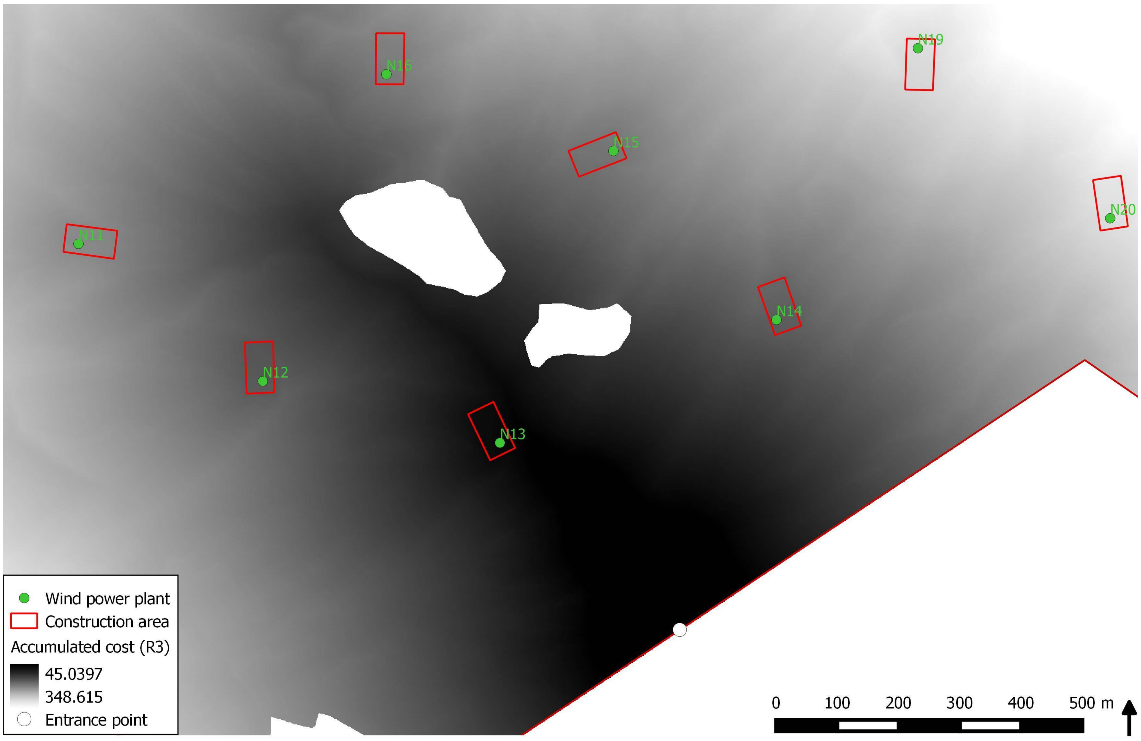


Figure 32. Scenario R3 accumulated cost surface in the central cluster.

In scenario R3, due to the structure of weighting, the accumulated cost surface heavily highlights the road heading northwest. The low movement costs around the road make it clearly visible even after passing the natural protection areas. The Boolean operation is also evident, with the natural protection areas having absolutely discrete borders.

Scenario G1

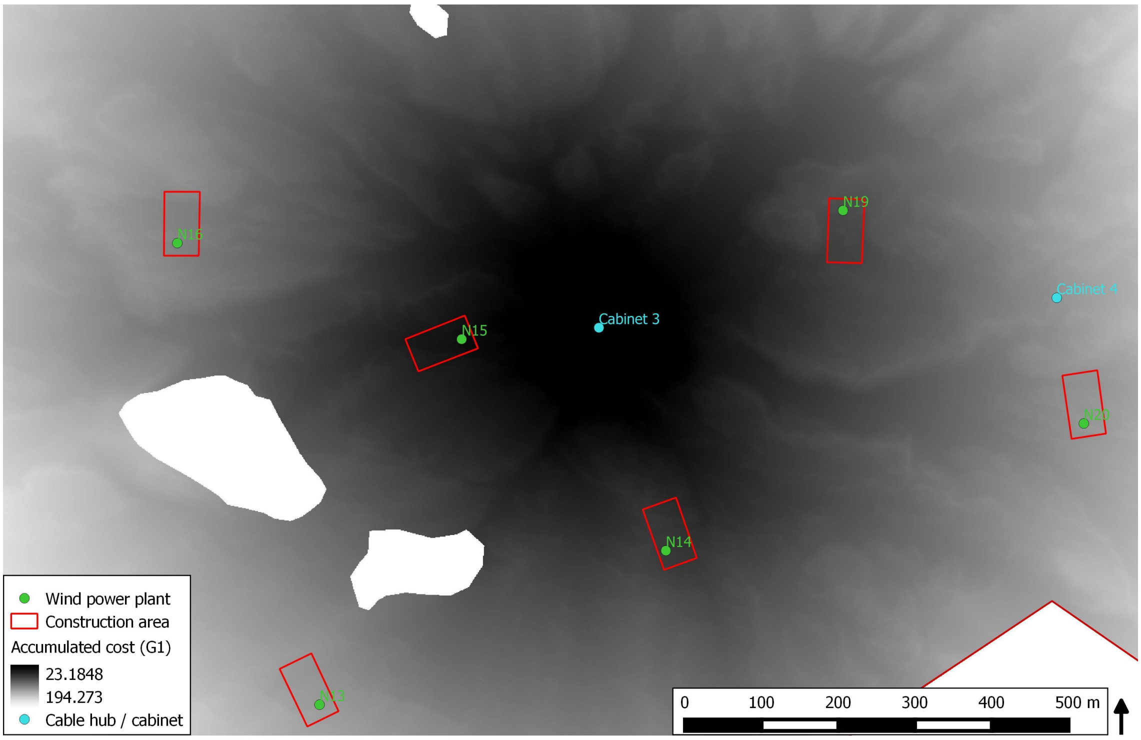


Figure 33. Scenario G1 accumulated cost surface with cable cabinet 3 as the starting point.

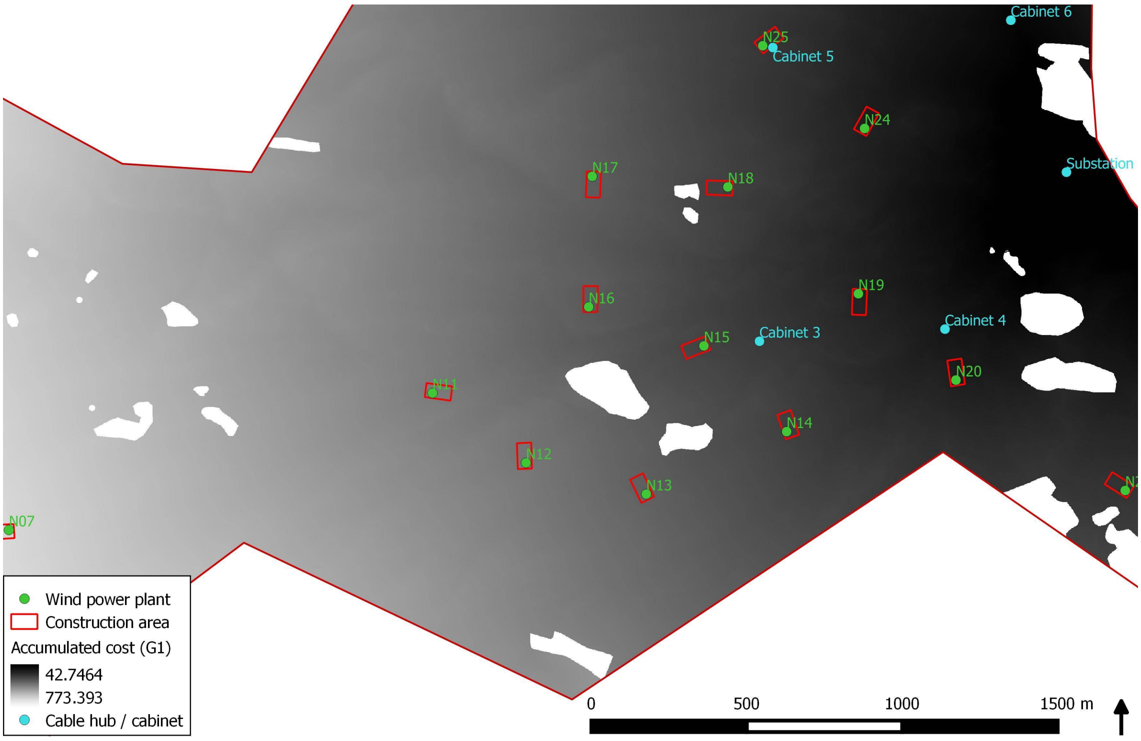


Figure 34. Scenario G1 accumulated cost surface with the substation as the starting point.

The spreading nature of the function is easily distinguished in the accumulated cost surface of G1 (Figure 23), as the starting point is not on the edge of project area, as in all of the road scenarios. The function disperses quite evenly in every direction, mostly following the patterns in slope gradient. In Figure 34, the accumulated cost surface starting from the substation visualizes the pathways of existing road network. The discrete polygons describing restriction areas demonstrates the Boolean nature of implementing them.

Scenario G2

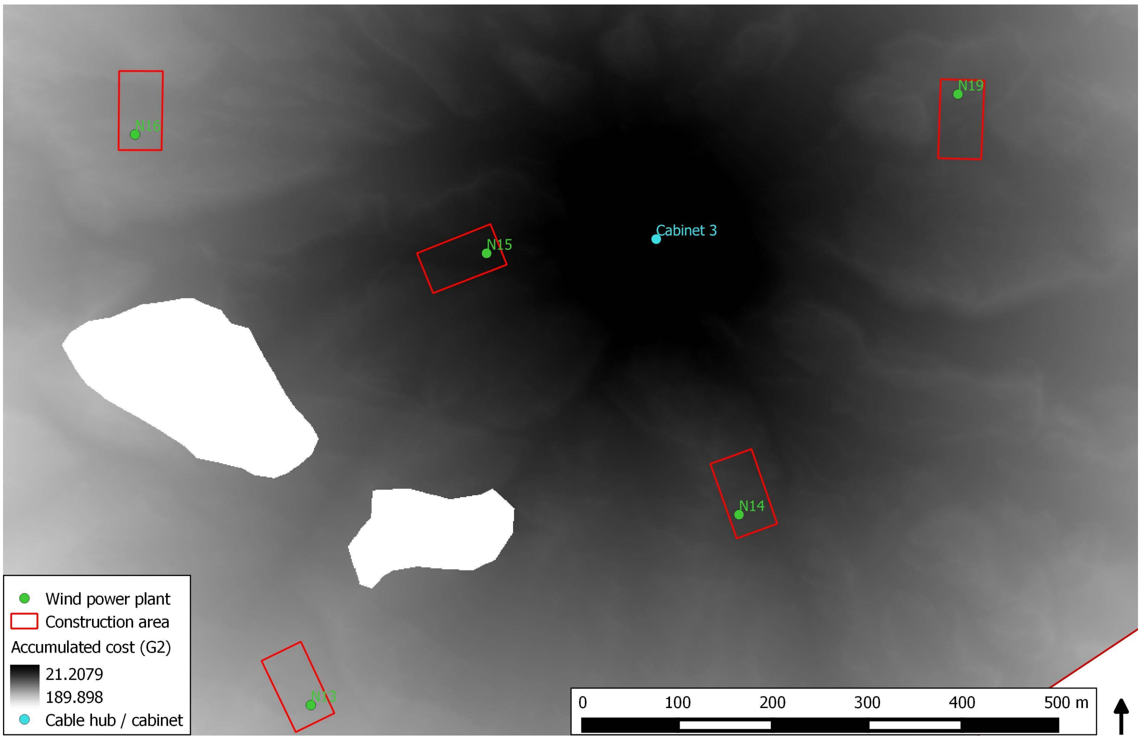


Figure 35. Scenario G2 accumulated cost surface with cable cabinet 3 as the starting point.

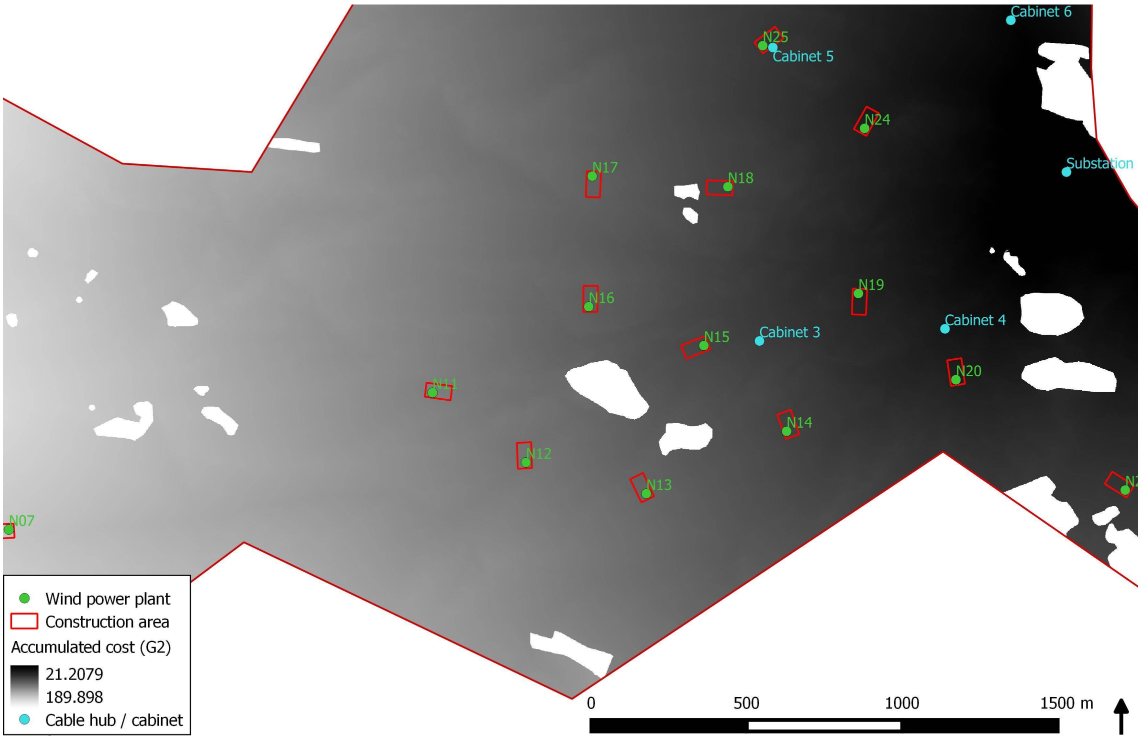


Figure 36. Scenario G2 accumulated cost surface with the substation as the starting point.

The accumulated cost surface of G2 is visually very similar to G1, with the exception that the result pathways of R3 are vaguely visible. The R3 road aligned between the natural protection areas towards power plant N15, and the long connection path that connects N07 with N12 are especially distinguishable.

4.4 Business Process Model and Notation (BPMN)

The main objective of the BPMN model produced for this study was to describe and document the activities, execution order of activities and the roles of different organizations in different phases of the overall process chronology. The documentation approach differs from the flowchart presented in Figure 14. The flowchart describes all the practical, technical and methodological aspects that are part of this study. The idea of the produced process model is to investigate this study as a single process in a business framework, in which other planning-related actors and activities are taken into account. This helps to understand the wider context of conducting a rather technical study inside a business.

Producing a process model also exposes the overall structure of the process to critical evaluation. This way different elements of the model can be monitored and further developed by using a standardized framework. Signavio Process Editor (version 9.4.1) using

BPMN 2.0 was used to create a documentation model of the LCP analysis process. It is a simple and easy-to-use tool and contains the core set of BPMN elements and objects.

The model produced for this study was designed to be clear and simple. Only the core elements of BPMN were used (Figure 37). Each organization have their own pools that graphically represent participants. Events, such as start and end events occur during the course of a process. Activities represent the work that is being performed. All the activities are connected and ordered by sequence flow lines. Message flows are also presented as lines, but they indicate the flow of messages between two participants. Gateways indicate some type of controlling of the process flow. A gateway with an X symbol represents a conditional XOR-gateway letting through only one of multiple possible sequence flows and a gateway with a + symbol represents an AND-gateway which is used for splitting and joining of sequence flow.

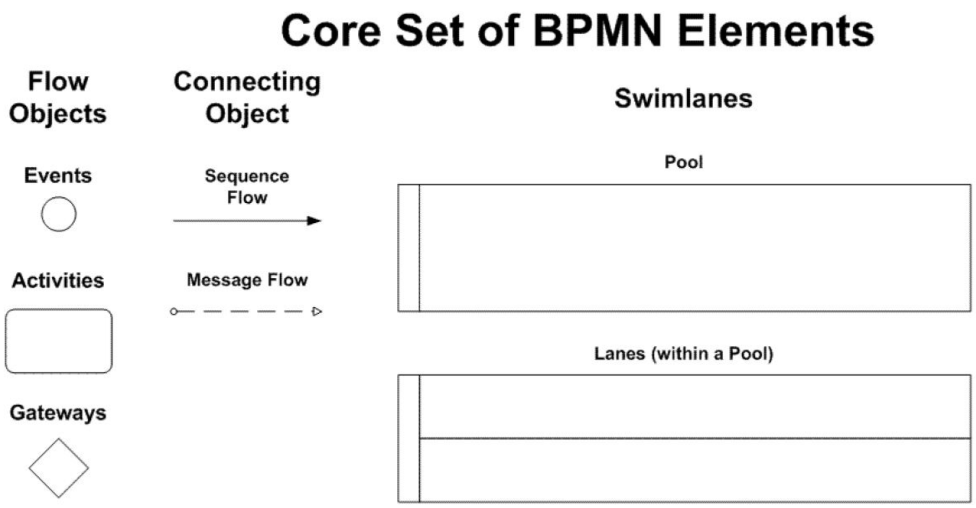


Figure 37. The core elements of a BPMN model (OMG 2011).

In general, the model has been kept simple. Every participant have their own sequence flows in the process, and message flows represent the information exchange between the organizations. One critical aspect in producing a BPMN model is the level of abstraction used. The modeler always has to search for a compromise, a sweet spot, between accuracy of details and the perceptibility of the model. Additionally, the modeler needs to determine, which aspects of the process are emphasized.

5 RESULTS

5.1 Infrastructure alignment

5.1.1 Access road network

The results of the study show that the produced least-cost paths correspond to the optimization objectives very well in different scenarios. R1 avoids slopes in a very aggressive manner, R2 circles around natural protection areas and tall vegetation while R3 has a more balanced approach as it is attracted by the existing road infrastructure. R3 also successfully evades all the restriction areas. Through these simple observations of the resulting least-cost paths, the optimization model can be validated. In other words, the produced least-cost paths represent optimal paths according to the assumptions and evaluations based on optimization objectives.

The results show that the optimization model produced road networks of very different lengths. The total length of road scenarios can be quantitatively compared to the preliminary road network presented in the master plan for Nordanå-Lövböle wind park (Table 4).

Table 4. Comparison table of road network lengths in different scenarios. Scenario lengths are compared to the master plan road network (Kemiönsaari 2014).

Scenario	**Length (m)**	**Difference (m)**
Master plan	24 068	0
R1a	28 937	+ 4 869
R1b	29 839	+ 5 771
R2	28 660	+ 4 592
R3	20 650	- 3 418
R3 (connected)	23 537	- 531

The scenario which was designed to minimize overall costs, R3, produced the shortest network. R3 is shorter in total length than the preliminary road network plan presented in the master plan. Even with the connection paths between different clusters included, R3 remains the shortest network overall. The preliminary road network of the master plan was structured around only one entrance point. It uses the same entrance point that was used in this study for the northern cluster.

Scenario R1

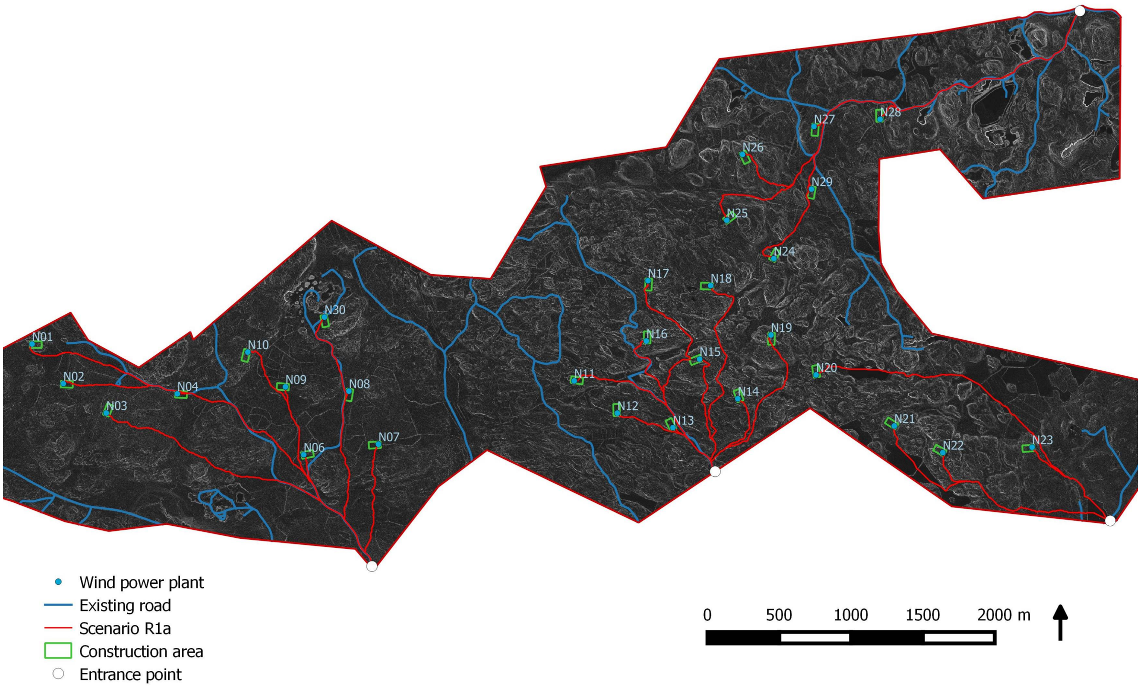


Figure 38. Scenario R1a LCP result paths on a slope map.

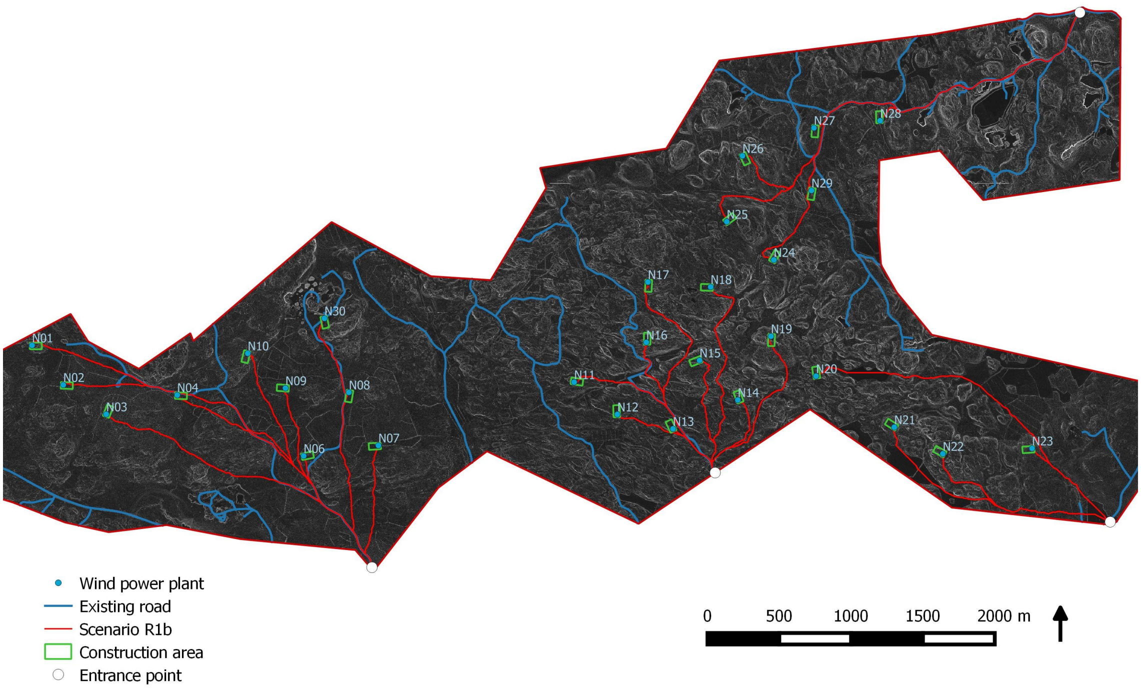


Figure 39. Scenario R1b LCP result paths on a slope map.

The results of scenario R1 represent the least-cost paths for minimizing movement in areas with high slope gradient. Of all the scenarios, R1 produced the longest road network. This is an expected result, because opting for routes with less slopes logically results in compromising the overall length. In real-life infrastructure planning, this leads to less groundworks and demolition works. The results should not be interpreted as a realistic scenario for infrastructure alignment. Instead it should be seen as a scenario that highlights the routes with low slope gradient.

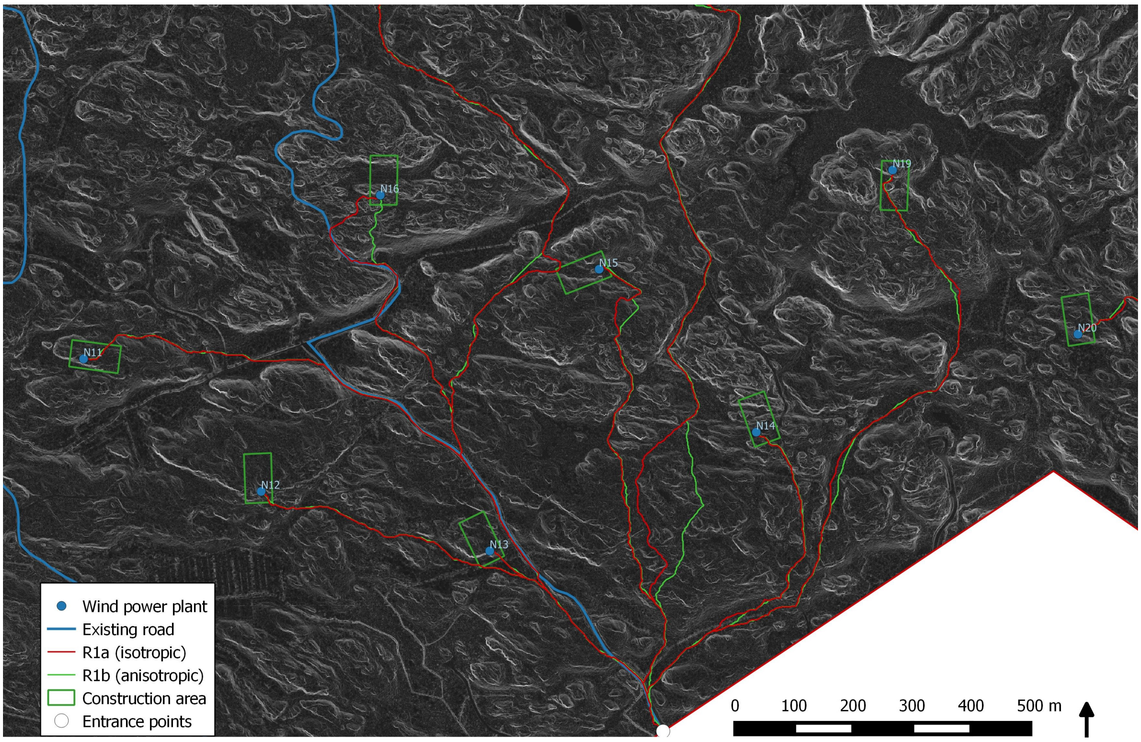


Figure 40. R1a and R1b central cluster result paths overlaid on a slope map.

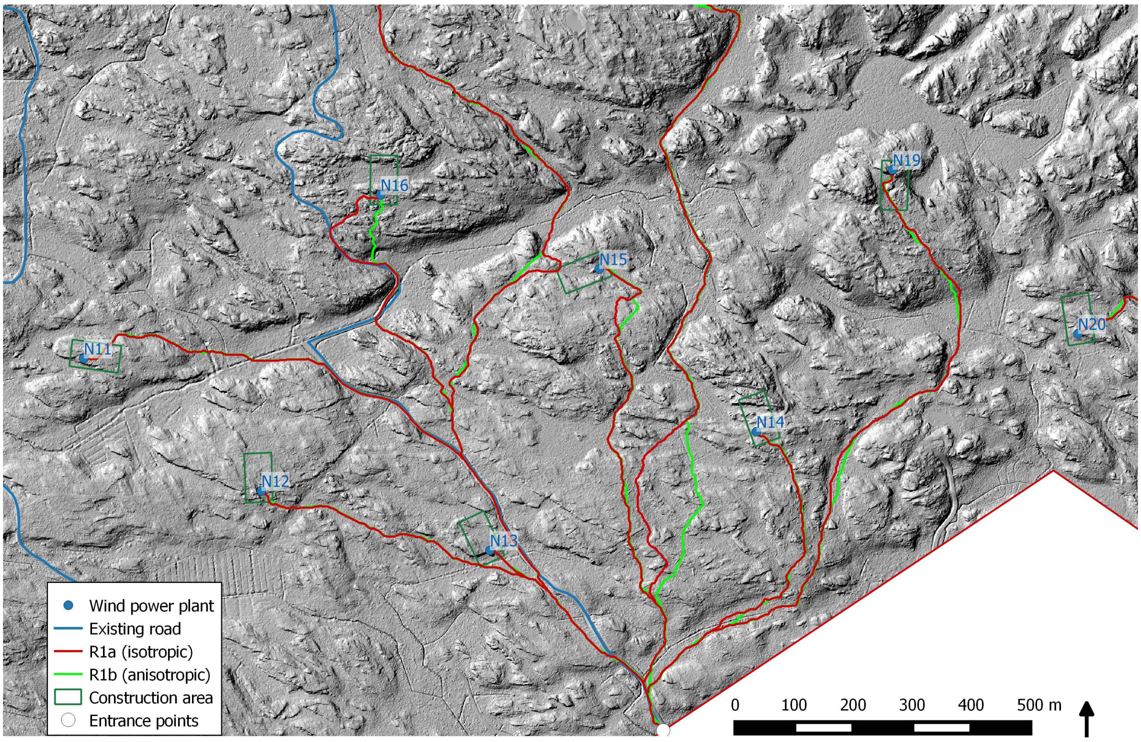


Figure 41. R1a and R1b central cluster result paths overlaid on a hillshade map.

When comparing the results of R1a and R1b, the anisotropic least-cost paths of R1b were moderately longer in overall length than isotropic least-cost paths. Still, the scenario R1 did not provide enough evidence to support the claims presented in literature that anisotropic spread produces more zigzagging or hairpin curves. However, by overlaying the results from both methods it is possible to see two different approaches to solving a problem, which might turn out to be valuable.

Scenario R2

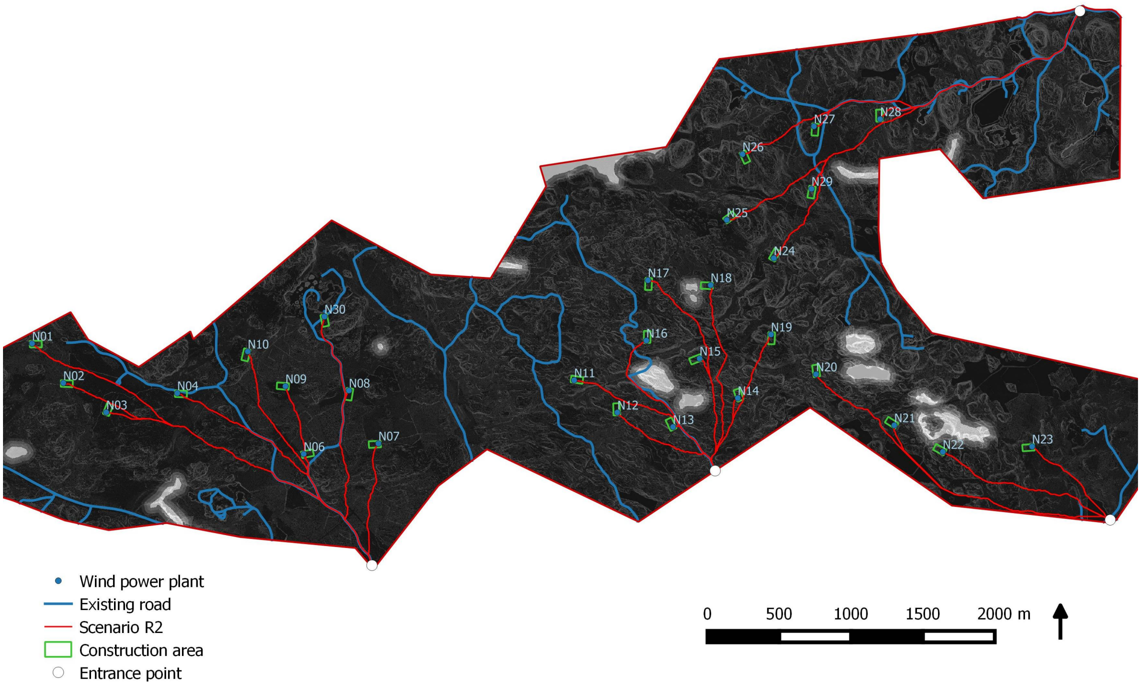


Figure 42. Scenario R2 LCP result paths on a cost-of-passage map.

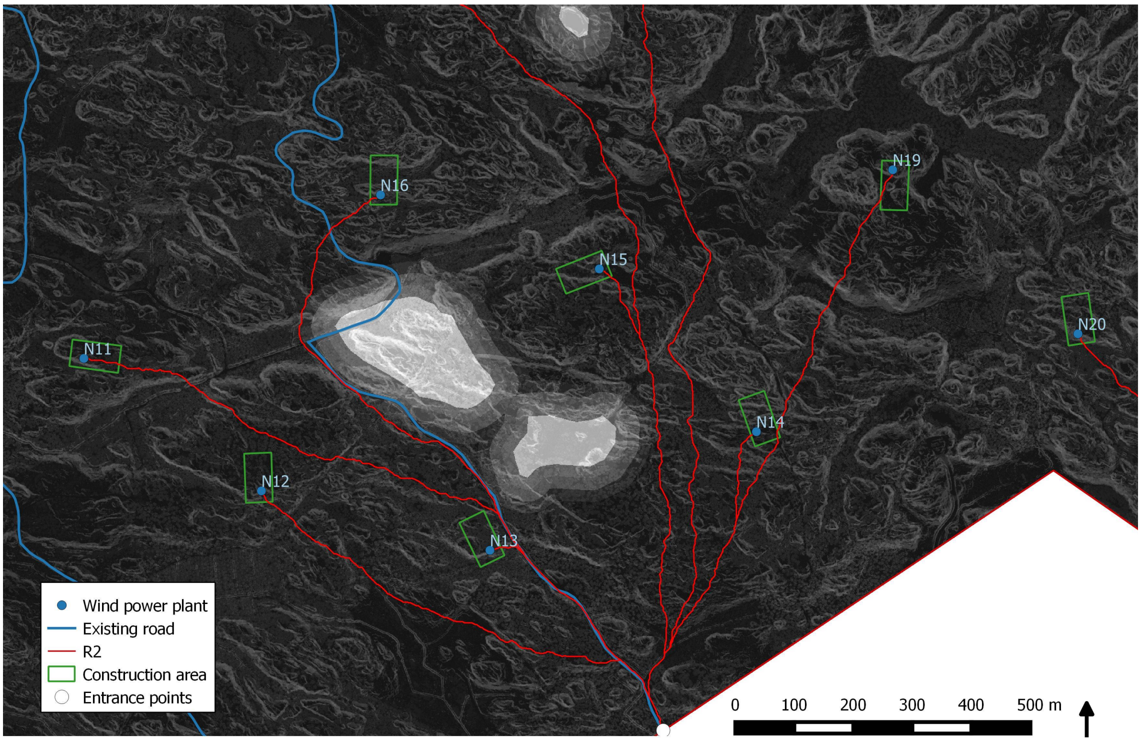


Figure 43. R2 central cluster result paths on a cost-of-passage map.

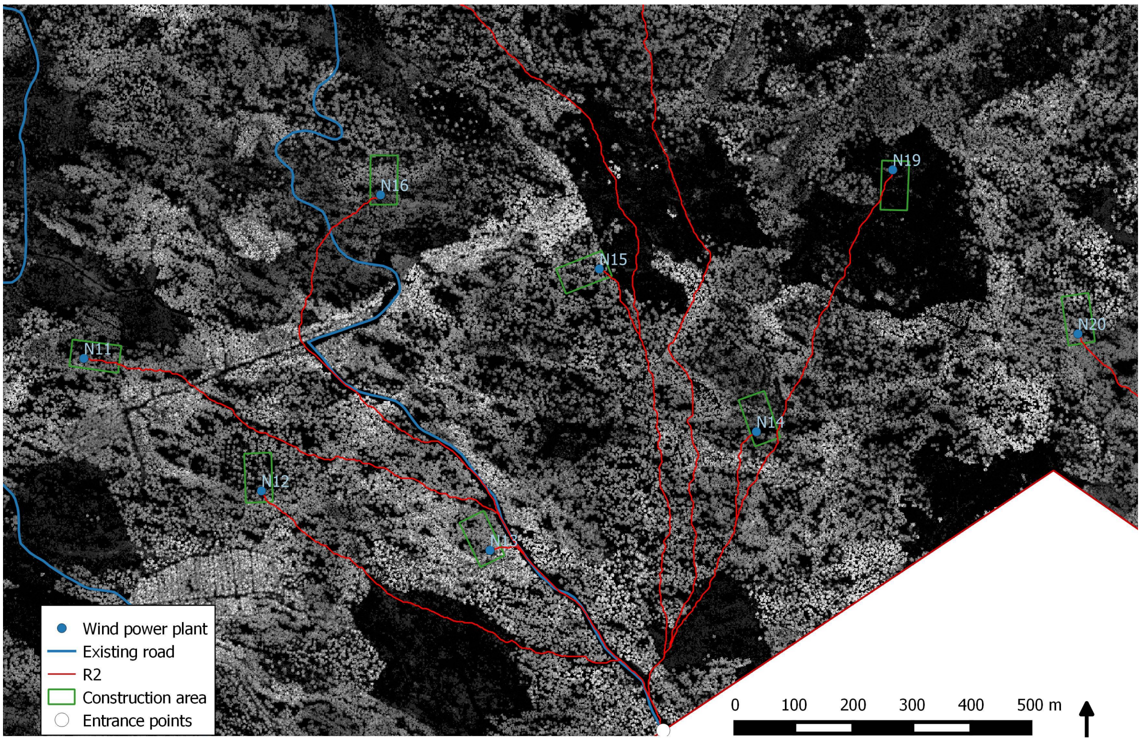


Figure 44. R2 central cluster result paths on a vegetation map.

The priority of R2 was to avoid natural protection areas and it succeeded in doing so. Otherwise the scenario produces pretty straightforward routes from the entrance points to the power plants. There is notably less zigzagging than in scenario R1, because the slopes are not as heavily weighted. Still, R2 avoids the steepest slopes and additionally, is clearly attracted by areas with low vegetation. For example, the path to N12 is notably affected by the vegetation factor, as the scenario takes route through a "bald spot" in vegetation.

Scenario R3

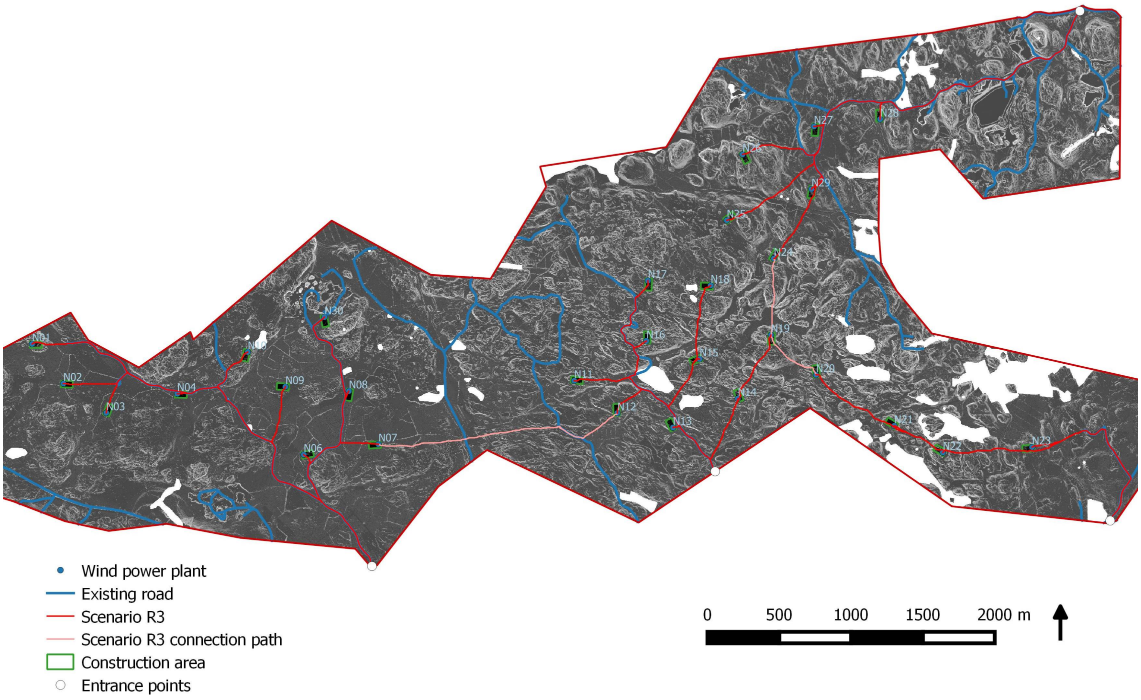


Figure 45. Scenario R3 LCP result paths on a cost-of-passage map.

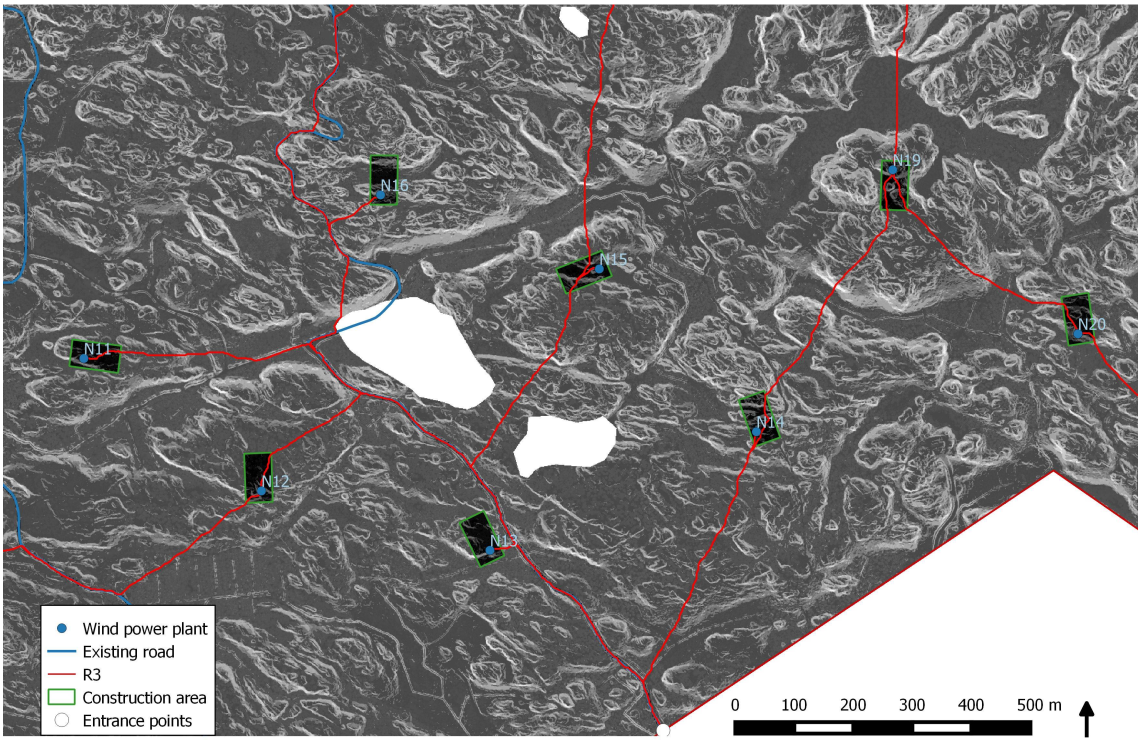


Figure 46. R3 central cluster result paths on a cost-of-passage map.

Scenario R3 was designed to find the least-cost paths by combining factors which have the most effect on the total infrastructure alignment costs. Thus, scenario R3 can be considered as the scenario which aims to minimize the costs. Construction areas and existing roads are clearly visible in the results as the areas with the lowest movement cost. Compared to R1, R3 compromises between the path length and the degree of slope gradient. In areas where R1 went around some slopes, R3 more straightforwardly calculates lower cumulative costs by just crossing some slopes. The connection paths to connect different clusters are presented in different color for easier comparison, because in other road network scenarios, the clusters were not connected. The connection paths were determined by intuition and other solutions were also in the consideration.

5.1.2 Internal electric grid

The results correspond to the optimization objectives regarding electric grid scenarios as well. Whereas G1 generate more off-road paths, G2 was provided with the optimized road network of scenario R3 to produce a higher level of road integration. Both grid scenarios also successfully evade all the restriction areas. According to the results, the optimization model produced cable networks of relatively similar lengths. The total length of grid scenarios can be quantitatively compared to the figures derived from a preliminary plan for the internal grid (Table 5). This preliminary plan is based on undisclosed offers that EFE has received from EPC contractors.

Table 5. Comparison table of grid lengths in different scenarios. Scenario lengths are compared to an indicative reference plan for internal grid. Two types of cables are used in the comparison: 150 mm and 300 mm.

Scenario	Length (m)	Difference (m)
Ref 150	15 000	0
Ref 300	19 000	0
Ref total	**34 000**	**0**
G1 150	20 892	+ 5 892
G1 300	13 349	- 5 651
G1 total	**34 241**	**+ 241**
G2 150	21 806	+ 6 806
G2 300	12 295	- 6 705
G2 total	**34 101**	**+ 101**

Both scenarios, G1 and G2, produced slightly longer overall networks than the reference plan. However, the main difference is that G1 and G2 use dramatically smaller amount 300 mm cables than the reference plan. It has to be pointed out, that in the reference plan, a different cabling structure was used. Compared to the six cable hubs that were used in G1 and G2 to connect circuits of turbines to the main substation, in the reference plan, eight substation connection hubs were used. This difference in the structure of the layouts partly explains the significant difference in the amount of 300 mm cables used.

Scenario G1

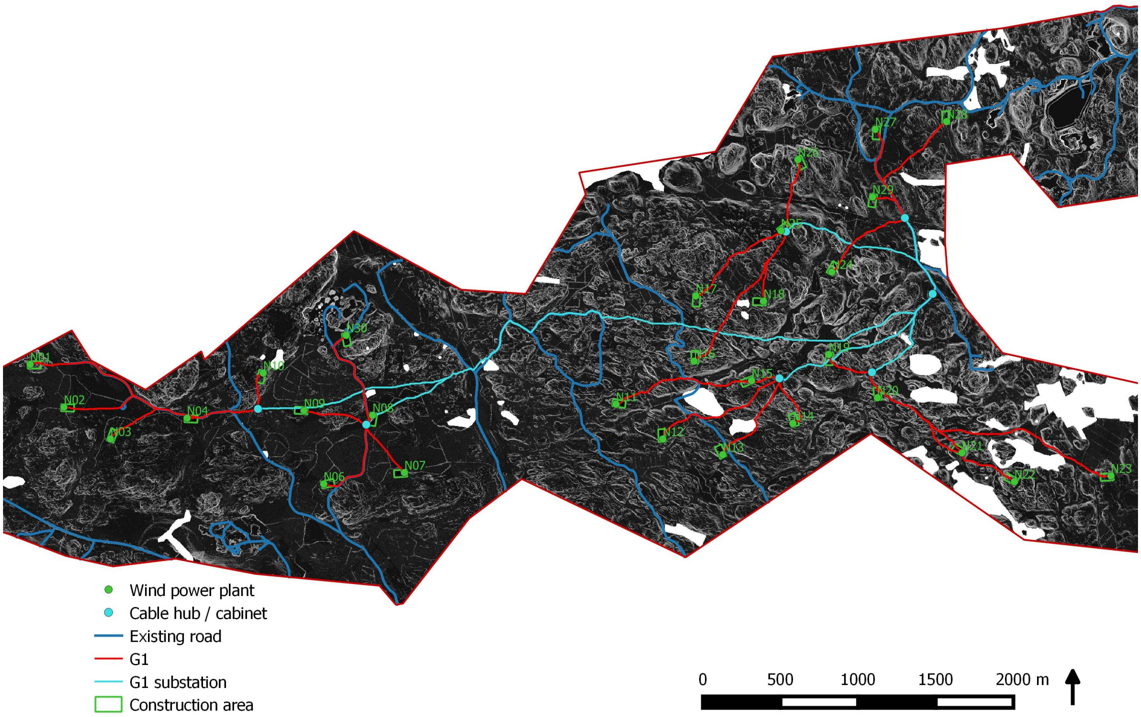


Figure 47. Scenario G1 LCP result paths on a cost-of-passage map.

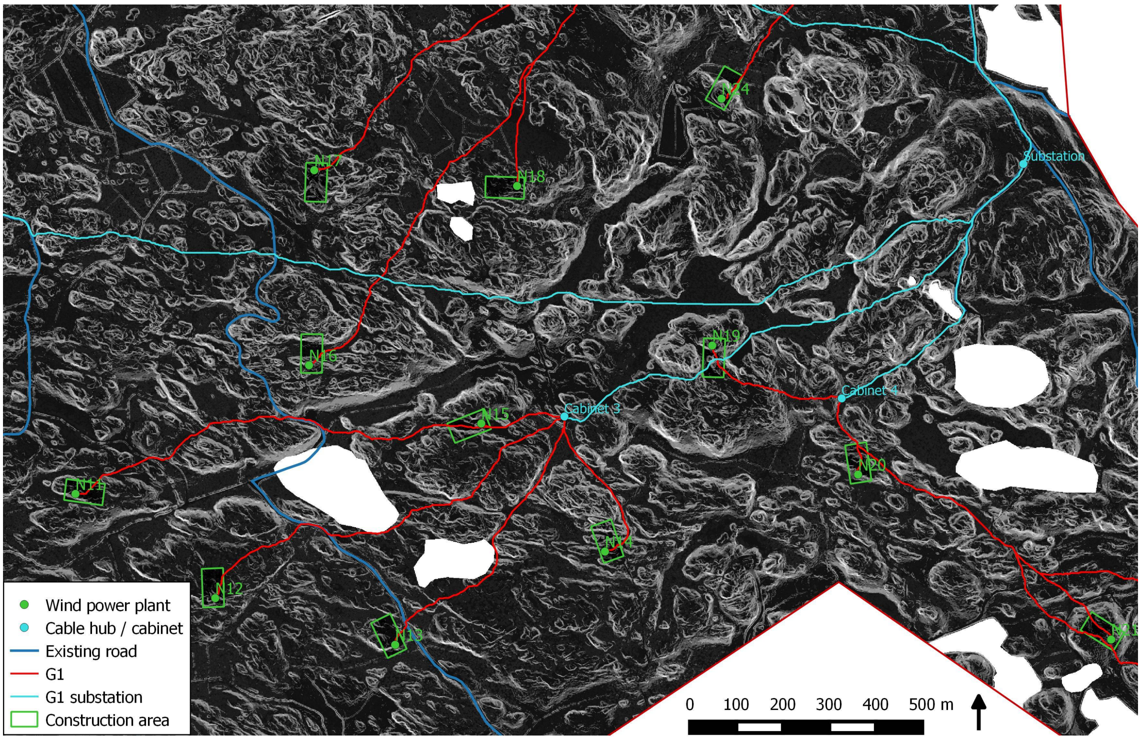


Figure 48. G1 central cluster result paths on a cost-of-passage map.

In scenario G1, the cabinets were first connected to their corresponding circuits of up to five power plants. The circuit structure was designed according to a preliminary plan by EFE Ab. After connecting the cable hubs to the turbines, all the cabinets were separately connected to the substation. Even though the existing roads were not as heavily weighted as in scenario R3, according to the result paths, the roads still proved out to be very attractive locations to place the cables. One example of this is the fact that the paths from cabinets 1 and 2 opted to take a longer route through north, where there are more existing roads available.

Scenario G2

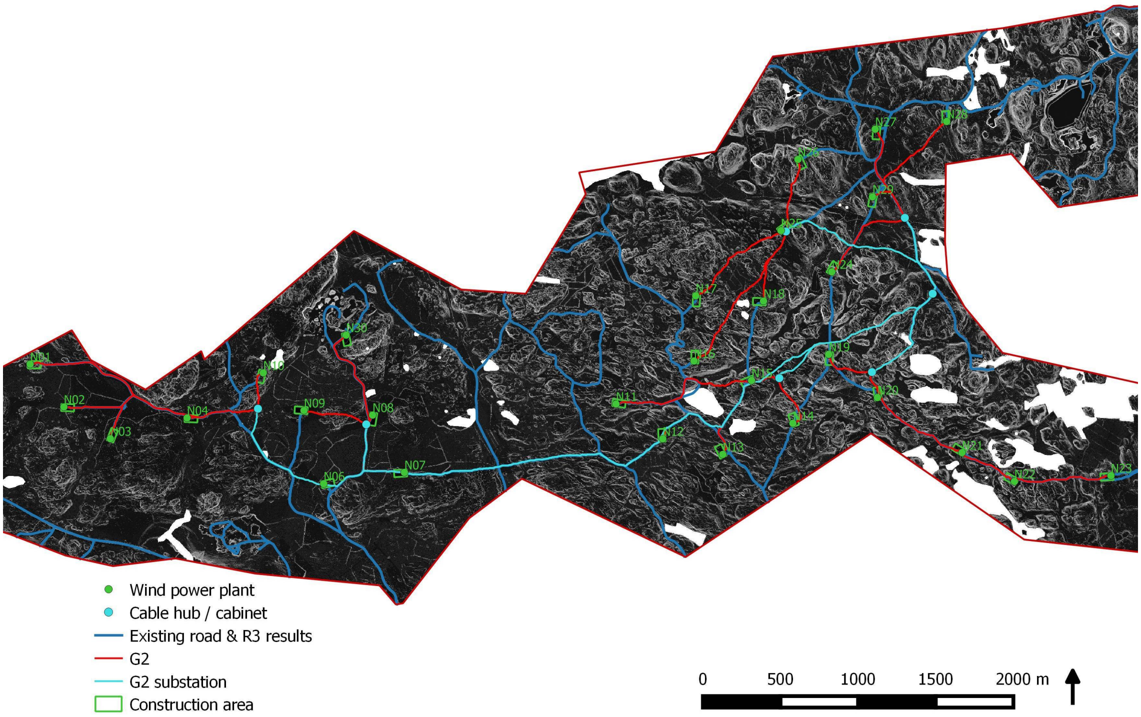


Figure 49. Scenario G2 LCP result paths on cost-of-passage map.

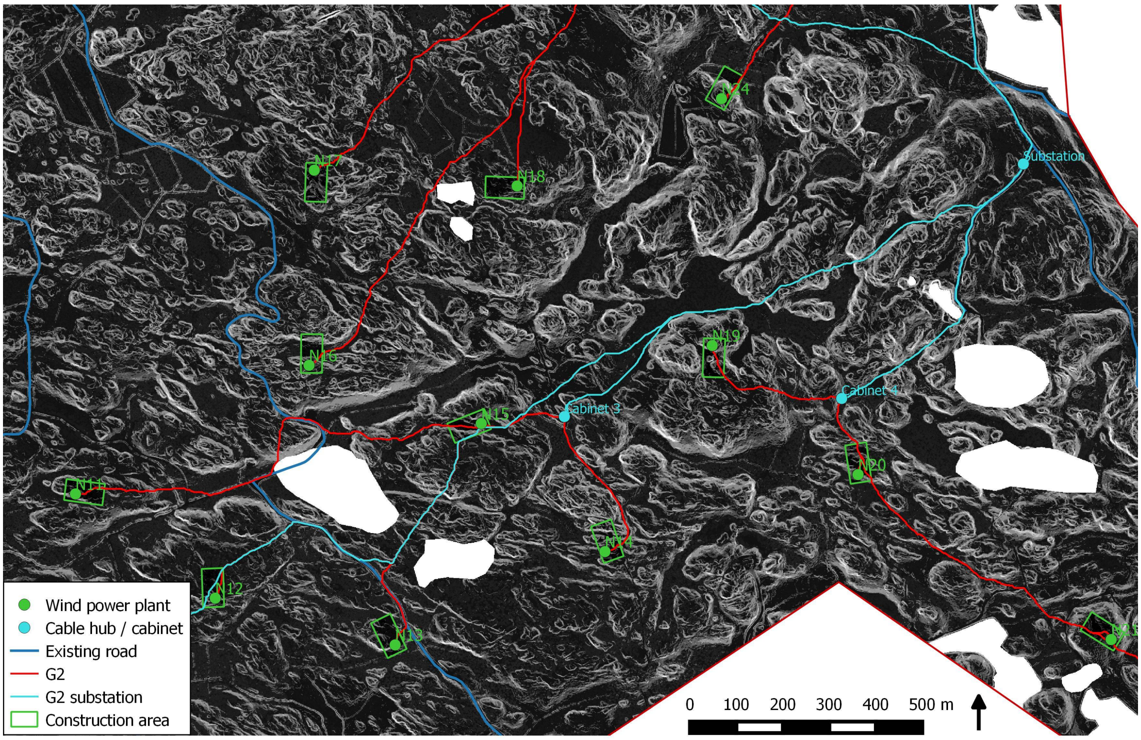


Figure 50. Scenario G2 LCP result paths on a cost-of-passage map.

LCP analysis result paths for scenario G2 were produced using the same parameters that were used in G1. Only difference with G2 was the cost-of-passage map which has result paths from scenario R3 integrated in the existing road network. This has a clear effect on the results: a lot of G2 paths follow the road paths of R3. The most evident results are in the eastern cluster and in the paths that connects cabinets 1 and 2 with the main substation. Also in the central cluster the grid paths of G2 switched to following R3 result paths.

5.2 Business Process Management Model

The result model of BPMN modeling is presented in Figure 51. There are five different organizations involved in the whole process and each one of them have their own pools. Every organization also have their own start and end events. *Decision maker* begins the process by ordering a LIDAR scanning for planning purposes from a *LIDAR supplier*. As the data is delivered, decision maker gives the data to *GIS Specialist* who processes the data, produces various different datasets of the data, designs factors and standardizes them for the analysis, designs scenarios, weights the criteria and combines the weighted factor maps. During this rather technical chapter that accounts for the methodological MCE phase of this study, the decision maker assists the GIS specialist with expert planning opinions and assessments. *Infrastructure company* begins its dedicated pool of activities as it is invited to a meeting with the decision maker to exchange information and assess the project. While the decision maker returns with this information to assist the GIS specialist, the infrastructure company consults the *turbine manufacturer* to gather information regarding wind turbine requirements and construction guidelines. They also discuss logistic options.

As GIS specialist finally runs the LCP algorithm with input data derived from the MCE-activities and exports the results, he or she presents the findings to the decision maker. Decision maker evaluates the results and requests more detailed offers for infrastructure alignment layout from the infrastructure company, who is closely working with the turbine manufacturer. Finally, the infrastructure company presents the decision maker with a layout proposal.

Two gateways are used in the model. When GIS specialist is processing lidar data and collecting additional data, such as planning data, an AND-gateway is used to split the sequence flow into two separate, but parallel operations. The second gateway is a conditional XOR-gateway, which determines the flow direction according to a defined condition. In this case, the potential use of a Boolean constraint is assessed. This phase continues to loop until all the scenarios have been processed through and the condition is met.

In the BPMN model produced in this study, a strong emphasize is on the communication and interactions between different organizations. For example, in a LCP analysis process, it may look evident that the GIS specialist who conducts the technical part of the study is the main actor in the LCP process from a technical perspective. However, it is essential that the specialist is provided with sufficient amount information and expert opinions during the process through communication. Therefore, even though the GIS specialist has more activities than other actors due to the technical emphasize of the model, other organizations are still as important as the GIS specialist for successful overall execution of the process.

The main emphasis in the model documenting the LCP process presented in this study, is in GIS specialist, who is responsible for carrying out the data processing, MCE and LCP phases. The decision maker orchestrates the work of GIS specialist and provides him or her data and expert information and finally evaluates the result and makes the final decision about the infrastructure alignment plan. As a typical outsourced service, the lidar supplier conducts remote sensing operations and then provides the main dataset for the decision maker. Infrastructure company and turbine manufacturer have small, but very important roles in being part of the process through expert consultation and definition of the practical framework for how infrastructure alignment planning should be implemented in wind power context.

The produced BPMN model represents one formal way of incorporating a technical analysis into a multi-organizational planning operation. The process approach gives not only a general level view on the workflow, but also a special insight on the relationship of internal and external activities in an infrastructure alignment planning process in wind power context.

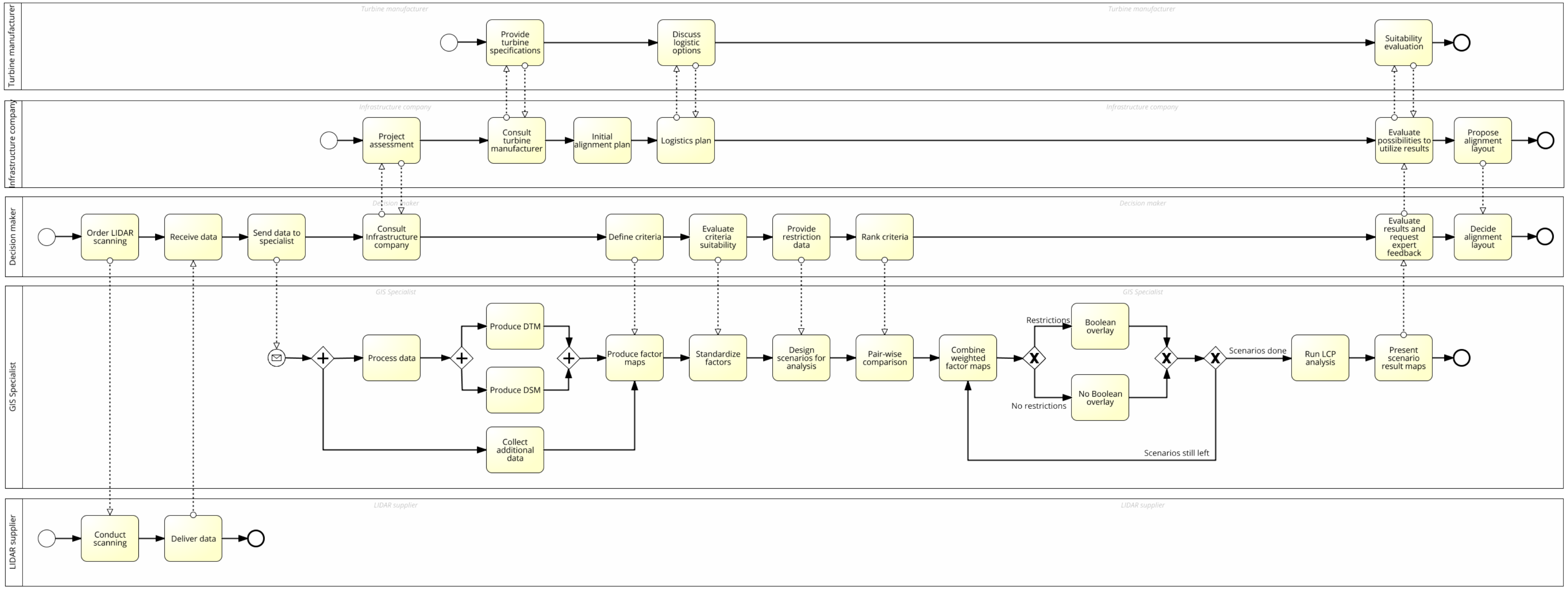


Figure 51. Least-cost path (LCP) analysis process modeled in BPMN 2.0.

6 DISCUSSION

6.1 Least-cost path modeling

This study has presented an extensive amount of least-cost paths to be utilized in onshore wind park infrastructure alignment. Input data for the analysis was derived from a high-resolution lidar data describing the level of terrain and surface elevation. There is a vast array of path optimization studies, which use LCP as their main research method in order to search for optimal locations for the construction of line-shaped infrastructure. However, this is the first study to utilize LCP specifically in wind park infrastructure alignment. Additionally, LCP literature generally consists of studies which use a maximum of 10 meter cell size in the analysis. The 1 meter cell size used in this study seems to be a truly outstanding resolution in the field of LCP studies. The lidar data would have enabled a cell size of 0.15 meters, but it is questionable how much this would contribute to the planning of paths that in reality are 6–8 meter wide. Additionally, even with a cell size of 0.5 meters, the calculation times increased significantly. After some initial simulations, resolution of 1 meter was determined as a good compromise between data accuracy and calculation times.

Antikainen (2013) argued that least-cost paths can only provide an approximation of the infinite number of possible movement options across a continuous surface that the raster surface is representing. He also suggested, that a relatively high resolution is the basis for producing even moderately accurate representation of the environment. In this study, the 1 meter resolution used is already heavily on the safe side of providing a sufficient resolution for accurate representation of the environment. This should be put once again in to the context of planning particularly wide roads.

In scenario R1, a comparison between isotropic and anisotropic spread functions was carried out. As there was not any significant differences in the result paths between these two functions in the study area, the isotropic spread was used in this study. This is due to the fact that the module that utilizes isotropic spread in GRASS GIS, *r.cost*, is more flexible than its anisotropic counterpart in defining the input data. As Jaga et al. (1993) have described, the absence of the directional component from the analysis is possible to compensate by applying higher friction values for steep slopes. In this study, this approach was present in the form of relatively high weights for slopes in each scenario.

In this study, the process of developing a wind park project and more specifically the alignment of infrastructure was used as a framework for the LCP process. Wind parks are generally planned in areas with sparse population and in Finland these areas are usually forested. LCP studies have traditionally optimized paths by only using physiographic datasets as input data. This thesis utilized a simple vector-to-raster conversion solution on how to include existing infrastructure, such as existing roads and preliminary construction areas in the LCP process.

It can be also pointed out, that even though having different factor maps for roads, slope and vegetation, in some degree they are overlapping in describing the real characteristics of the environment. For example, even if using the slope gradient as the only factor in a scenario, similarly to R1, the existing roads are indirectly included in the cost-of-passage map through the fact that roads are generally flat and tend to have low slope gradients. The tendency of the LCP result paths to follow the roads at least in some degree, is clearly evident. From vegetation perspective, there were some isolated patches of high vegetation among otherwise sparsely vegetated area. In some cases, these may represent small hills, which have been left untouched during extensive forest clearings.

In some scenarios, such as R3, the existing roads are heavily favored in weighting. This choice can be supported with the simple fact that an existing road generally decreases the road construction costs to half. In addition to doing so, heavily weighting existing roads clearly bundles together multiple roads to different turbines. This is useful when trying to optimize the total length of a network of least-cost paths, because the algorithm works in a way that calculates routes from a starting point to stopping points completely separately. This means, that even when executing mass operations with the algorithm, the result is not a network, it is merely a group of independent paths that are overlaid and merged. In other words, the paths are “blind” to each other.

The results of the LCP analysis were compared to the preliminary planned roads presented in the 2014 master plan of the wind park project. One of the objectives of this study was to present optimal paths for wind power infrastructure alignment according to different optimization objectives. The objective was not to minimize particularly the total length of infrastructure networks, but to minimize the total accumulated impedance to movement in various optimization scenarios. Nevertheless, the total length of a network still remains a valid way to quantify and compare the results. In scenario R3, the aim was to minimize expensive excavations and demolition works, to inclusively utilize existing road infrastructure and conservatively avoid areas with high vegetation. Even so, R3 turned out to produce a shorter road network than the preliminary plan. Also, while the reference plan has only one logistical entrance point to the project area, R3 provides four separate entrance points. Scenarios R1 and R2 produced relatively long network of roads. While they may not necessarily represent the most realistic overall solutions for infrastructure alignment, these networks correspond to the optimization objectives.

The grid length results show that the differences between G1, G2 and the reference plan were very small. When comparing the result lengths with the reference plan lengths, it has to be pointed out, that in the preliminary reference plan, circuits of 2–4 turbines were designed and connected to the substation. Also, on the contrary to the layout produced in this study, no particular cable hubs were used. Instead, one turbine in a circuit is connected to the main substation. In the reference plan, a total of eight circuits of turbines are connected to the main substation. This means that the reference plan has a structure of eight connection points to the substation, whereas scenarios G1 and G2 have only six points. This structural difference is the main reason for the fact, that G1 and G2 use dramatically less 300 mm cables than the reference plan. When evaluating raw cable costs as an investment, this is a crucial notion, as 300 mm cables are generally larger, have more conducting capacity and consequently are more expensive than 150 mm cables.

The obtained results give a strong indication of the possibilities of using carefully evaluated algorithm-based path optimization process as the basis for infrastructure alignment planning.

The objective of the LCP analysis was to find optimal paths. As Longley et al. (2005) put it, optimality in path planning means minimizing a variable that is related to the optimization objective. In this study, multiple different optimization objectives were defined. The factors behind these targets did not represent any exact, quantifiable figures of monetary costs or losses to protected natural habitats. Instead, it has to be emphasized, that the factors behind optimization objectives were based on expert assumptions of which features of the environment have the most effect on the overall costs or the balance of ecosystems. It can be argued, that extensive interviews in the criteria standardization and criteria weighting phases could have increased the credibility of the results. However, it was not within the scope of this thesis to organize comprehensive interviews, as the study emphasis was more on the technical side of applying LCP in GIS platform.

One suggestion for future research would be to study the possibilities of modifying the algorithm to produce least-cost paths iteratively. As mentioned in the previous chapter, when executing mass operations, the separate paths are blind to each other. In the case of including existing road network to the analysis, it could provide much more realistic results if every segment of produced path would be integrated into the input data to be utilized by the calculation of next segment. This type of iterative approach could develop the path incrementally, segment-by-segment. If LCP analysis would be developed further into the direction of an iterative approach, one of the first challenges would be to solve the constant rewriting processes involved with large raster datasets. Even during this study, the editing of high resolution large raster datasets proved out to be quite time- and resource-consuming process.

6.2 Methodological considerations

Establishing a logical methodology that combines various different methods utilizing different types of data is a challenging process that requires careful evaluation. As Malczewski (1999) suggested, the balanced combining of hard and soft information is integral to producing relevant information for a decision making process. In this study, both information types, hard and soft, were utilized and carefully combined in order to successfully carry out the optimization process.

The presented methodology consisted of multiple established methods. However, there were some choices made in the selection of single methods, especially in the multi-criteria evaluation phase. Traditionally, the method of weighted linear combination (WLC) has been used throughout a MCE process from standardization of factors to combining the factors by weighted averaging (Jiang & Eastman 2000). Instead of using WLC, the standardization of different factors was conducted in this study by using fuzzy measures due to the strong evidence on its benefits (Bonham-Carter 1994; Eastman 1999; Jiang & Eastman 2000; Atkinson et al. 2005; Choi et al. 2009; Zimmermann 2010).

Despite being a very essential transformation phase regarding the study results, and despite conducting the transformations by using widely accepted and validated methods, there could be an endless debate about the validity of the subjective assumptions that were used to determine fuzzy values and scenario weights. However, acknowledging the nature of the MCE process, multiple scenarios were used to compensate its subjectivity. Therefore, it can be argued that the results of this study reflect a number of options for different optimization goals and objectives. Additionally, subjective judgement is a part of the fundamental nature of the fuzzy set theory the fuzzy logic is based on (Bonham-Carter 1994). Furthermore, it was not within the scope of this study to aim for an absolutely precise quantification of cost factors in wind power infrastructure construction. Firstly, it would be very difficult due to the very dynamic and complex nature of infrastructure pricing and secondly, one of the main objectives of this study was to present a new applied methodological approach.

Nevertheless, more studies for determining exact definitions of wind power infrastructure cost factors can be suggested, in order to deepen the understanding of overall cost structures involved in the planning process. Additionally, if there ever were exact, generally validated figures available for different cost factors, they could be implemented in the LCP approach using various different methods, including multi-criteria evaluation.

The methodology of this thesis was also documented from a process approach using BPMN as the method. In the model, the execution of the LCP analysis phase itself was not described in detail in this model. This decision was made because there are multiple different methods and software to execute the calculation of an accumulated cost surface. However, it could be argued, that a separate subprocess could have been be produced out of the LCP analysis -activity in the model. The subprocess would have used a low-level abstraction to describe specific elements of the analysis, such as different modules and parameters used. Other potential subprocesses to include some external stakeholders could have been created as well. For example the interaction between the decision maker and land owners, or between decision maker and the public officers of the municipality could have been described.

The model was abstracted so that it only describes the essential activities of the LCP process. In reality, in the technical part of the process, a countless number of editing and conversion operations was conducted. Additionally, the communication between different organizations was only roughly estimated. In reality, the communication between turbine manufacturer, infrastructure company, decision maker and the GIS specialist is much more fluid. The technical part of the model was also strongly emphasized, as data acquisition and MCE processes were documented accurately

BPMN model approaches the study from a broader perspective, as a business process in planning context. It aims to give a quick organizational insight into how LCP-based infrastructure alignment planning could be implemented to a business framework. The strongest argument for using BPMN was the standardized nature and technical versatility of the tool. In addition to being a method for documentation of process activities and interrelationships, process modeling is also used for development of processes. These features give standardized process models extra depth compared to traditional, generic flowcharts. Having originated in the field of information systems and profoundly applied in information technology, process modeling has matured in an environment where agile and flexible development of systems and processes has been the ideal.

In the context of this thesis, in addition to clarifying and documenting the research process, the produced BPMN model should indicate a signal of continuation in the scientific research of the LCP process. As Campagna et al. (2014) suggested, a process model can either be descriptive (as-is) or prescriptive (to-be). The BPMN produced in this study is descriptive, as it aims to present a new method to be used in the field of infrastructure alignment planning. Future research could suggest new ways (to-be) of arranging activities between organizations to further develop the efficient execution of LCP process through the concepts of BPMN.

In Chapter 2.6.2, the pioneer work of Campagna et al. (2014) around the concept of metaplanning was presented. Metaplanning applies BPM and BPMN to land use planning. The work of Campagna et al. is a strong indication that there is potential in incorporating land use planning and process modeling. The BPMN model presented in this study continues the work of Campagna et al. in the sense of importing a technical GIS-based process into the framework of land use planning.

The LCP analysis process presented in this study was theoretically approached from the aspects of GIS and environmental modeling. During the analysis, remotely sensed terrain data, infrastructure expertise and an algorithm applied for environmental modeling purposes were imported to a GIS platform to produce a model for searching patterns in the environment in order to find least-cost paths. There are some theoretical challenges in producing such a model containing different types of data elements.

As Harvey (1969) argued, the problem and the power of computer syntaxes lies in their ability to link abstract linguistic elements to real world phenomena. The *r.cost* module is built around the Dijkstra's algorithm, but as there is no exact definition of what type of data the input data can contain, certain transformations were conducted in order to assign such values to the input data, that correspond real world phenomena. The linguistic link between abstract computer languages and intelligible human language was achieved by manipulating the input data by using fuzzy logic to reflect human assumptions and objectives, such as criteria for road construction. As Zimmermann (2010) argued, fuzzy set theory can act as a bridge between human language and formal models.

Kemp (1992) also underlined the importance of creating interfaces between reality and spatial data in order to produce real representations of the environment. The method for standardization used in this study was based on fuzzy logic, and it can be seen as the bridge between raw spatial data and reality in terms of human objectives and practices. Zimmermann (2010) referred to fuzzy set theory as "computational intelligence". In this study, fuzzy logic provided a framework that represents suitability assumptions for certain human operations, such as construction of roads. In order to use this framework, raw spatial data and human assumptions were combined. Within this methodology, the raw spatial data values were transformed into fuzzy values that represent new contextual information.

As Martin (1991) has suggested, if a cartographical model represents the relationship between a map and a reality, a map can be seen as a representation of reality. Applying this analogy to the results of this study, with expert assumptions representing the relationship between an algorithm and terrain data, the results of this study represent a model or as Harvey (1969) put it, a formalized expression of theory, of optimal paths in real-world terrain. Similarly to Chorley's (1964) suggestion of mapping a part of the reality into the model in order for the model to become a theory of the real world, the lidar data represents a part of the reality in this model.

Martin (1991) also presented a transformation-based view of GIS operation (Figure 1). In this study, data is not only manipulated in T3 of Martin's concept, but also an algorithm is implemented on the processed data to produce completely new data which is based on the input data. Martin also argued that the transformations that are carried out in a GIS operation are the single most important factor on the output data quality. Although in his chart (Figure 1), Martin seemingly discussed transformations conducted solely for hard, quantitative data, in this study, also soft, qualitative data was integrated into the data model. The most challenging transformation operations conducted in this study were the subjective assumptions used in the standardization and the weighting of the factors.

Malczewski (1999) suggested that GIS is a process which aims to support decision making by refining raw data into useful information. This very well concludes the ultimate purpose of the model presented in this study. Raw datasets of different types were combined, processed and refined into geographical information. This data is intended to support decision making. These objectives are also supported by Brimicombe's (2010) *solution space* which urged for co-operation between the utilitarianism of environmental engineering, model simulation and GIS. As the BPMN model produced in this study proposes, engineers of an infrastructure company should certainly have a role in evaluating the utilitarianism and realism of the suggested results from simulation model produced with the tools and methods of GIS.

The philosophy that was the catalyst behind this study can be very well described with the characteristics of the new paradigm of environmental modeling. The phrase Haggett & Chorley used almost 50 years ago would be a surprisingly accurate way to describe the objectives and results of this study. Even though Dijkstra's algorithm is even older than Haggett's & Chorley's study, applying the algorithm in the field of geography and more specifically using it to search real world terrain data, is a typical example of the new models and formulas which create "new patterns of searching the real world". With the level of resolution offered by modern lidar scannings, the phrase could be reformulated for this study to "searching for micro-level environmental patterns of the real world".

Antikainen (2008) has claimed that the subject of path planning and path optimization has not been widely studied in academic geography. This is mainly because the concepts and theories of path-finding are based on graph theory, mathematics and computer science. Nevertheless, Antikainen has called for bringing the topic of terrain-based path planning back into geography, where he feels it truly belongs to. He also emphasizes that GIS as a science is supported by a decent amount of theoretical knowledge, particularly in the field of spatiality.

6.3 Future applicability

There is no evident reason why the presented LCP process could not be applied to other infrastructure-related issues as well. With the resolution that modern lidar scannings offer, even the optimization of very narrow trails would be possible if sufficient processing capacity was available. It can be argued that the level of remote sensing data that is available today is not very well understood in the infrastructure planning sector. The relevance for this study arose from the traditional and rather conservative general practices that infrastructure companies in Finland are using. As presented in Chapter 2.5.2, infrastructure planning is generally conducted by manual drawing operations by using thematic maps, in some cases even in paper format. In a time, which is heavily labeled by the revolution of digital information, it seems that there truly is a technological blind spot in infrastructure planning.

This study was carried out by using mostly open data, with lidar data being the only exception. This makes the study more transparent and easier to repeat. While already available in some parts of Finland, open lidar data of the National Land Survey of Finland (NLS) could be utilized in conducting a similar analysis in many other potential study areas.

In terms of generalizing the results of this study, the study area should be interpreted in multiple clusters due to the varying construction factors inside the project area as discussed in Chapter 3. The four cluster system used in the LCP analysis could be perceived as forming four separate wind parks. As the western cluster is topographically the easiest regarding groundworks, the central, eastern and northern clusters are much closer and thus easier to connect with the main electrical substation. As the central, eastern and northern clusters all require a large amount of new roads to be built, in the eastern cluster it is possible to connect the roads and cables pretty straightforwardly, as the turbines are geometrically quite aligned. These examples indicate that in the context of infrastructure construction, even though Nordanå-Lövböle might be more challenging than the average wind park projects already in operation, there is still a lot of variation inside the project. This notion can be used to suggest that by separately evaluating the results of each cluster, the generalization possibilities are more evident.

As this study was carried out in the context of wind power infrastructure requirements, it is relevant to identify the industry-specific specifications that were not yet included in this study and could be considered for future research. Data that could be more extensively utilized for LCP analysis in the context of wind power, is road requirements. In order to assess more in detail the suitability of planned paths for the transportation of wind power plants, the road requirements specified by wind turbine manufacturers could be further included in the path planning algorithm. These road requirements include specific minimum dimensions for width, flatness, carrying capacity, curve angle, curve radius and slope gradient of the roads. Slope gradient was the first requirement that was truly implemented in this study as a part of the factor standardization phase.

According to Yu et al. (2003), in order to prevent the used path optimization algorithm from producing very sharp curvature, a simple directional component could be added to the syntax of the algorithm. This way the road requirements defined by a turbine manufacturer could be much more accurately included in the analysis. Additionally, Goncalves (2010) presented an approach to allow the modeling of wide paths with a fixed width to be used in cases where the raster cost surface has a higher resolution than the width of the planned path. These procedures would require more specialized expertise in programming and thus they were not within the scope of this study.

As mentioned before, the results of this LCP process should be primarily seen as supportive information for planning of infrastructure alignment. Although there were some phases in the process that were in some degree exposed to subjectivity, the LCP process should provide a significantly more objective method for analyzing patterns of the environment than the traditional way of carrying out field observations. As Goodchild (1991:336) described the possibilities of GIS, perhaps the results of this study can "provoke profound geographical thoughts" in the fields of wind power and infrastructure alignment planning. After all, infrastructure alignment should not be seen only as an engineering issue as which it is usually perceived, but also as a geographical, spatial research topic. Bringing the topic back to geography, as Antikainen (2008) suggested, could truly refresh the philosophical foundation of the approaches and practices used in the planning of infrastructure alignment. Geography, and more specifically GIS, can provide a new framework for the preliminary mapping and understanding of the features of the environment that goes beyond the capabilities of human perception. Combined with the human interpretation and expertise on environmental phenomena, this could provide a platform for a new paradigm in the planning of infrastructure alignment.

ACKNOWLEDGEMENTS

This study would not have been conducted if it was not for Ansgar Hahn, the CEO of EFE Ab, the co-operation company that admitted me the scholarship to carry out this study. Ansgar Hahn is a visionary project manager, who has a passion for developing the field of wind power project development. During this study, he continued to provide me with invaluable expert knowledge and insights regarding different aspects of wind power project planning. I am grateful for his trust in me to develop a new, geographical approach for solving a problem, which is traditionally addressed by more pragmatic sciences.

I would also like to thank Adjunct Professor Niina Käyhkö, who as my instructor contributed greatly in helping me to organize my draft papers into a more articulated and logically structured end result.

LITERATURE

Anderegg, W.R.L, J.W. Prall, J. Harold, S.H. Schneider (2009). Expert credibility in climate change. Proceeding of the National Academy of Sciences of the United States of America, 107: 27, 12107–12109.

Anderson, A.E. & J. Nelson (2004). Projecting vector-based road networks with a shortest path algorithm. Canadian Journal of Forest Research, 34: 7, 1444–1457.

Antikainen, H. (2008). Developing methods for GIS-based terrain path planning. Nordia Geographical Publications. 37: 6, 63–72.

Antikainen, H. (2013). Using the Hierarchical Pathfinding A* Algorithm in GIS to Find Paths through Rasters with Nonuniform Traversal Cost. International Journal of Geo-Information. 2: 4, 996–1014.

ArcGIS Resource Center, (2012). LAS Dataset To Raster (Conversion). 11.11.2015. < http://resources.arcgis.com/EN/HELP/MAIN/10.1/index.html#//001200000052000000>

Atkinson DM., P. Deadman, D. Dudycha, S. Traynor (2005). Multi-criteria evaluation and least cost path analysis for an arctic all-weather road. Applied Geography, 25: 4, 287–307.

Bannon, L. J. (1995). The Politics of Design: Representing Work. Communications of ACM, 38: 9, 66–68.

Berzan, C., K. Veeramachaneni, J. McDermott (2011). Algorithms for Cable Network Design on Large-scale Wind Farms. MSRP Technical Report, Massachusetts Institute of Technology.

Bonham-Carter, G. (1994). Geographic Information Systems for Geoscientists: Modelling with GIS. 1st ed. 398 p. Elsevier.

Brimicombe, A. (2010). GIS, Environmental Modeling and Engineering. 2. p. 380 s. Taylor & Francis Group.

Burgess, R. & C. Darken (2004). Realistic Human Path Planning Using Fluid Simulation. Proceedings of the Behavior Representation in Modeling and Simulation Conference.

Burrough, P. (1992). Are GIS data structures too simple minded? Computers and Geosciences. 18: 4, 395–400.

Campagna, M., K. Ivanov, P. Massa (2014). Orchestrating the spatial planning process: from Business Process Management to 2nd generation Planning Support Systems. Proceedings of the AGILE’2014 International Conference on Geographic Information Science.

Choi, Y., H.D. Park, C. Sunwoo, K.C. Clarke (2009). Multi-criteria evaluation and least-cost path analysis for optimal haulage routing of dump trucks in large scale open-pit mines. International Journal of Geographical Information Science. 23: 12, 1541-1567.

Chorley, R.J. (1964). Geography and Analogue Theory. Annals of the Association of American Geographers. 54: 1, 127–137.

Chou, Y.-H. (1997). Exploring Spatial Analysis in Geographic Information Systems. 500 s. Onword_Press, Santa Fe, USA

Collischonn, W., J.V. Pilar (2000). A direction dependent least-cost-path algorithm for roads and canals.
International Journal of Geographical Information Science. 14: 4, 397–406.

Cormen, T., C. Leiserson & R. Rivest (1997). *Introduction to Algorithms.* 2. p. 984 s. The MIT Press. Cambridge, Massachusetts.

Coulter, E.D., J. Coakley, J. Sessions (2006). The Analytic Hierarchy Process: A Tutorial for Use in Prioritizing Forest Road Investments to Minimize Environmental Effects.
International Journal of Forest Engineering 17: 2, 51–69

Dijkstra, E.W. (1959). A note on two problems in connexion with graphs.
Numerische Mathematik 1, 269–271

Empower (2014). A powerpoint presentation by Petri Koski: *Considerations in wind power plant 110 kV grid connection planning*. 35 s.

Energy Department (2010). A report by Empower. *Sähköverkkokomponenttien yksikkö-hintojen määritys.* 39 s. 15.10.2015.
<https://www.energiavirasto.fi/documents/10179/0/Lahde_9_Empower_Yksikkohinnat_2010+%282%29.pdf/4844805e-7bd2-460b-8e96-9dcc11186262>

Eastman, J.R. (1999) Multi-criteria evaluation and GIS.
In Longley, Paul A., Michael F. Goodchild, David J. Maguire, and David W. Rhind, (eds.): *Geographical Information Systems: Principles, Techniques, Management and Applications. Second Edition.* Chichester, UK Wiley.

EFE (2012). Kemiönsaaren Nordanå-Lövbölen tuulivoimapuisto – ympäristövaikutusten arviointiselostus. 2.2.2016.
<http://www.kemionsaari.fi/media/tekniska/pdf/planer/aktuella/generalplaner/lovbole_vindkraft/forslag/%C2%A7121-3-4_18.6.2014_liite2b_n-l_yva-selostus.pdf>

EFE (2015a). Varsinais-Suomen Energia Oy – The Project. 8.12.2015.
<http://www.efe.fi/?page_id=153>

EFE (2015b). "Logistic map of Nordanå-Lövböle wind park project."
Image from the internal planning material of Varsinais-Suomen Energia Oy.

EPO (2015a). "An example of a wind park collector system."
Image from internal planning material of energy without pollution GmbH.

EPO (2015b). "A four-turbine circuit connected to PCC through a cable hub."
Image from internal planning material of energy without pollution GmbH.

EWEA (2009). *Wind Energy – The Facts. A guide to the technology, economics and future of wind power.* 1st ed. 592 p. Earthscan, London.

Finnish Energy (2016). Energiavuosi 2015 – Sähkö. 20.1.2016.
<http://energia.fi/kalvosarjat/energiavuosi-2015-sahko>

FTA (2016). Finnish Transport Agency: Digiroad – A National Road and Street Database. 1.2.2016. <http://www.digiroad.fi/en_GB/>

Gasch, R. & J. Twele (2012). Wind Power Plants. Fundamentals, Design, Construction and Operation. 2nd ed. 567 p. Springer, Berlin Heidelberg.

Gietl, R., M. Doneus and M. Fera (2008). Cost Distance Analysis in an Alpine Environment: Comparison of Different Cost Surface Modules.
In: Posluschny, A., K. Lambers and I. Herzog (eds.): *Layers of Perception. Proceedings of the 35th International Conference on Computer Applications and Quantitative Methods in Archaeology (CAA)*

Goodchild, M. F. (1991). Just the facts.
Political Geography Quarterly 10: 4, 335–337.
Goodchild, M. F. (1992). The State of GIS for Environmental Problem-Solving
The State of GIS 8–15
Goodchild, M. F. (2009). Geographic information systems and science: today and tomorrow. *Procedia Earth and Planetary Science* 1, 1037–1043.
GRASS GIS manual (2012). Raster commands: *r.drain*. 1.2.2016.
<https://grass.osgeo.org/grass64/manuals/r.drain.html>
GRASS GIS manual (2015). Raster commands: *r.walk*. 1.2.2016.
<https://grass.osgeo.org/grass64/manuals/r.walk.html>
IEA (2013). Technology Roadmap – Wind energy, 2013 edition. 8.12.2015.
<http://www.iea.org/publications/freepublications/publication/Wind_2013_Roadmap.pdf>
IEA Wind (2015). 2014 Annual Report. 8.12.2015.
<http://www.ieawind.org/annual_reports_PDF/2014/2014%20AR_smallfile.pdf>
IPCC (2014). Schlömer S., T. Bruckner, L. Fulton, E. Hertwich, A. McKinnon, D. Perczyk, J. Roy, R. Schaeffer, R. Sims, P. Smith, and R. Wiser: Annex III: Technology-specific cost and performance parameters.
In: Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel and J.C. Minx (eds.): *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.
Haggett, P., & R.J. Chorley (1967). Models, paradigms and the new geography.
In: Chorley, R.J. & P. Haggett, P. (eds.): *Models in Geography*, Methuen, 19–41.
Harvey, D. (1969). *Explanation in Geography* 521 s. Edward Arnold Ltd, London.
Herzog, I. (2014). Least-cost Paths – Some Methodological Issues.
Internet Archaeology 36
Jaga, R. M. N., A. Sundaram, T. Natarajan (1993). Wasteland development using geographic information system techniques.
International Journal of Remote Sensing, 14: 17, 3249–3527.
Jiang, H., R. Eastman (2000). Application of fuzzy measures in multi-criteria evaluation in GIS.
International Journal of Geographical Information Science, 14: 2, 173–184.
Kaldellis, J. K. & D. Zafirakis (2011). The wind energy (r)evolution: A short review of a long history.
Renewable Energy 36, 1887–1901.
Kemiönsaari (2014). Yleiskaavat, Nordanå-Lövbölen tuulipuiston osayleiskaava. 18.1.2016. <http://www.kemionsaari.fi/media/tekniska/pdf/planer/aktuella/generalplaner/lovbole_vindkraft/ehdotus_ii_nordanal%C3%B6vb%C3%B6le_tuulipuiston_oyk_18-6-2014_%C3%A4ndring_16-9-2014.pdf>
Kemp, K. K. (1992). Environmental Modeling with GIS: A Strategy for Dealing with Spatial Continuity. 142 s.
National Center for Geographic Information and Analysis. Technical Report 93-3.

Klap, A., Saaristo, H., Juvonen, T. (2011) Varsinais-Suomen tuulivoimaselvitys 2010–2011, 100 s. 1.12.2015. <http://www.varsinais-suomi.fi/images/tiedostot/Maankaytto/2011/Tuulivoima/tuulivoimaselvitys2010_2011.pdf>

Lee, J., & D. Stucky (1998). On applying viewshed analysis for determining least-cost paths on digital elevation models. International Journal of Geographical Information Systems, 12: 8, 891–905.

Lillesand, T.M., R.W. Kiefer, J.W. Chipman (2008). Remote Sensing and Image Interpretation. 6. P. 765 s. John Wiley & Sons.

Liu, K., J. Sessions (1993). Preliminary planning of road systems using digital terrain models. International Journal of Forest Engineering, *4*: 27–32.

Longley, P.A., M.F. Goodchild, D.J. Maguire, D.W. Rhind (2011). Geographic Information Systems & Science. 3rd ed. 560 p. John Wiley & Sons.

Lynn, P.A. (2012). Onshore and Offshore Wind Energy: An Introduction. 1st ed. 240 p. John Wiley & Sons.

Maguire, D.J., M.F. Goodchild & D.W. Rhind. (1991). Geographical Information Systems Vol. 1. 447 s. Longman Scientific and Technical, London.

Mainguenaud, M. (1995). Modeling the Network Component of Geographical Information Systems. International Journal of Geographical Information Science 9: 6, 575–593.

Malczewski, J. (1999). GIS and Multicriteria Decision Analysis. 392 s. John Wiley & Sons.

Malczewski, J. (2000). On the Use of Weighted Linear Combination Method in GIS: Common and Best Practice Approaches. Transactions in GIS 4: 1, 5–22.

Martin, D. (1991). Geographic Information Systems and Their Socioeconomic Applications. 182 s. Andover, Hants, England: Routledge, Capman, and Hall.

OMG Object Management Group, (2011). Business Process Model and Notation (BPMN), version 2.0, 1.12.2015. <http://www.omg.org/spec/BPMN/2.0/PDF>

Paikkatietohakemisto (2015). Metatieto. Korkeusmalli 10m. 1.2.2016. <http://www.paikkatietohakemisto.fi/catalogue/ui/metadata.html?uuid=cd640425-315f-4b12-86c5-192e98701dcb&lang=fi>

Paikkatietohakemisto (2016). Metatieto. Laserkeilausaineisto. 1.2.2016. <http://www.paikkatietohakemisto.fi/catalogue/ui/metadata.html?uuid=0e55977c-00c9-4c46-9c87-dee6b27d2d5c&lang=fi>

Picard, N., L. Gazull, V. Freycon (2006). Finding optimal routes for harvesting tree access. International Journal of Forest Engineering, *17*: 35–48.

Pickles, J. (1995). Representations in an Electronic Age: Geography, GIS, and Democracy. In Pickles, J. (ed.): Ground Truth: The Social Implications of Geographic Information Systems, 637–663.

Rees, W. G. (2004). Least-cost paths in mountainous terrain. Computers & Geosciences, 30: 3, 203–209.

Rogers, L. (2001). PEGGER & ROADVIEW – A New GIS Tool To Assist Engineers in Operations Planning.

The International Mountain Logging and 11th Pacific Northwest Skyline Symposium, Seattle, University of Washington: 177–182.

Rogers, L. (2005). Automating contour-based route projection for preliminary forest road designs using GIS. MSc Thesis. 87 p. College of Forest Resources, University of Washington.

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology* 15: 3, 234–281

Sheppard, E. (1995). GIS and society: towards a research agenda *Cartography and Geographical Information Systems* 22: 1, 5–16

Sheppard, E. (2000). Quantitative geography: representations, practices, and possibilities. *Environment and Planning D: Society and Space* 19: 5, 535–55

Smith, H. & P. Fingar (2003). *Business Process Management: The Third Wave*. 1. p. 311s. Meghan-Kiffer Press.

STY (2016). Suomen Tuulivoimayhdistys, Suomen tuulivoimalaitokset- ja hankkeet, 20.1.2016. <http://www.tuulivoimayhdistys.fi/hankelista>

STY (2015). Suomen Tuulivoimayhdistys, Tuulipuiston rakentaminen ja voimaloiden pystytys, 21.10.2015. <http://www.tuulivoimayhdistys.fi/tietoa-tuulivoimasta/tietoa-tuulivoimasta/tuulivoimaprojekti/puiston-rakentaminen-ja-voimaloiden-pystytys>

Taylor, P. J. (1990). GKS. *Political Geography Quarterly* 9: 3, 211–212

Tuulivoimaopas (2014). Tuulivoima Suomessa, 21.10.2015. <http://www.tuulivoimaopas.fi/tuulivoima_suomessa>

U.S. Department of Energy (2015). Wind manufacturing and supply chain, Infrastructure and Logistics. 11.1.2016 <http://energy.gov/eere/wind/infrastructure-and-logistics>

Vaasa Energy Institute (2010). Tuulivoiman kustannukset. 12.10.2015. <http://vindkraft.fi/public/index.php?cmd=smarty&id=37_lfi>

Van Leusen, M. (2002). Pattern to Process. Methodological Investigations into the Formation and Interpretation of Spatial Patterns in Archaeological Landscapes. PhD thesis 356 p. Rijksuniversiteit Groningen.

Waters, N. (2002). Modeling the Environment with GIS: A Historical Perspective from Geography. *In* Clarke, K. (ed.): *Geographic Information Systems and Environmental Modeling.* 1–35.

Weske, M. (2012). *Business Process Management.* 2nd ed. 403 p. Springer Berlin Heidelberg.

Wizelius, T. (2007). *Developing Wind Power Projects. Theory and Practice*. 1st ed. 296 p. Earthscan Publications, London.

Wright, D., M. Goodchild & J. Proctor (1997). GIS: Tool or Science? Demystifying the Persistent Ambiguity of GIS as “Tool” versus “Science”. *Annals of the Association of American Geographers* 87: 2, 346–362.

Yu, C., J. Lee, M. J. Munro-Stasiuk (2003). Extensions to least-cost path algorithms for roadway planning. *Geographical Information Science* 17: 4, 361–376

Zimmermann, H.-J. (2010). Fuzzy set theory. *WIREs Comp Stat* 2: 3, 317–332.