Nordia
Geographical Publications

Volume 41:4

Accessibility, population change and scale dependency

Exploring geospatial patterns in Finland, 1880–2009

Ossi Kotavaara

ACADEMIC DISSERTATION
to be presented with the permission of the Doctoral Training Committee for Human Sciences of the University of Oulu Graduate School (UniOGS), for public discussion in the lecture hall L10, on 5th of December, 2012, at 12 noon.

Nordia
Geographical Publications

Volume 41:4

Accessibility, population change and scale dependency

Exploring geospatial patterns in Finland, 1880–2009

Ossi Kotavaara

Nordia Geographical Publications

Publications of
The Geographical Society of Northern Finland
and
The Department of Geography, University of Oulu

Address: Department of Geography
P.O. Box 3000
FIN-90014 University of Oulu
FINLAND
juho.luukkonen@oulu.fi

Editor: Teijo Klemettilä

Nordia Geographical Publications
ISBN 978-952-62-0035-4
ISSN 1238-2086

Juvenes Print

Oulu 2012

Accessibility, population change and scale dependency

Exploring geospatial patterns in Finland, 1880–2009

Contents

Abstract

Accessibility, population change and scale dependency – Exploring geospatial patterns in Finland, 1880–2009

Keywords: Accessibility, population change, scale, geographical information system (GIS), geospatial, generalized additive models (GAM), Finland

The aim of this thesis is to study the effect of transport accessibility on population change in Finland and the matter of scale to this relationship. During the research period, 1880−2009, Finland has urbanized, changed from an agrarian to a post-industrial economy and the transport system has developed from being manually powered to one that is fully motorized. The theoretical framework of the study is founded on transport and economic geography and regional science. The analytical basis is in the geospatial analysis consisting of geographic information systems (GIS), which is applied to accessibility analyses and data management, and generalized additive models (GAM), multiple regression enabling non-linear relationships, used in exploring the statistical relationships. The study utilizes precise GIS data from transport networks and the population having both a high spatial accuracy. To achieve a long temporal reach in analysis, data for historical population and transport networks were digitized. The eight scales of the analysis begin from municipal and built-up area scales, continuing to grid cells at six resolutions, reaching the resolution of 2×2 km. Accessibility indicators applied in the study are based on fastest route calculations. In addition to the road-based potential accessibility, also rail and air transport accessibilities are considered as travel times to nearest facilities. In addition, analysing the effect of accessibilities on population change, also the effect of key socio-economic variables and population density are tested, to avoid omitted variable problems and enhance the performance of models.

The key finding of the thesis is that road network-based potential accessibility captured the attractiveness of cities and population concentrations well, in general, since the end of the 19th century. Particularly at accurate resolutions, the explanatory power of the models strongly relies on potential accessibility, which is computed with steep distance decay function. While accessibility has a concentrating effect on activities on a regional scale, it also has a decentralising effect within urban areas. After modelling the major population trend, concentration at centre-edge-periphery axis, the effect of other variables could be estimated. The accessibility of railways had a slight effect on population change between 1920−1980. Airport accessibility included in the analysis since 1990 and air transports seem to have an attracting effect on the population. By including key socio-economic variables into models, the explanatory and predictive performance of the models increased. The major trend prevailing in models explaining population change at different scales is that the explanatory power and predictive performance decreases with the scale that is becoming more accurate. Thus, it can be concluded that the correctness of the presumptions of scale are important in analysing the effects of accessibility. As population change could be modelled on a grid cell basis by using

accessibility and socio-economic variables, the model or its sub-components can be applied to make more spatially detailed estimations of population change. The models of the study show also that non-linear relations between accessibility and population change are very common.

The results of this thesis thus outline the status quo of Finnish population trends. Spatially and temporally consistent models show that over the last two decades, the population has concentrated towards easily accessible areas at a regional scale, but population pressure exists at urban margins at a local scale. Nothing indicates that the trend would not continue and the growth of the peripheries in general does not seem to be realistic in the future. As an analytical framework, the significance of this thesis lies less in defining direct causalities, rather exploring and revealing spatial trends and patterns that may be hidden to intuitive thinking and cartographical examination. The results of the study are country specific, but the study setting and the modelling procedure can be applied wherever enough accurate GIS data of transport infrastructure, population and also socio-economic conditions are available.

Supervisors

Professor Jarmo Rusanen
Department of Geography
University of Oulu
Finland

Professor Miska Luoto
Department of Geosciences and Geography
University of Helsinki
Finland

Post-doctoral Researcher Harri Antikainen
Department of Geography
University of Oulu
Finland

Pre-examiners

Professor Markku Tykkyläinen
Department of Geographical and Historical Studies
University of Eastern Finland
Finland

Post-doctoral Researcher Katarina Haugen
Department of Geography and Economic History
Umeå University
Sweden

Official opponent

Senior Lecturer Tuuli Toivonen
Department of Geosciences and Geography
University of Helsinki
Finland

List of original papers

This thesis consists of a synopsis and four research articles, which are referred to by their Roman numerals in the text.

Article I Kotavaara, O., H. Antikainen & J. Rusanen (2011). Urbanization and Transportation in Finland, 1880-1970: A GIS Analysis. *Journal of Interdisciplinary History* 42: 1, 89–109.[1]

Article II Kotavaara, O., H. Antikainen & J. Rusanen (2011). Population change and accessibility by road and rail networks: GIS and statistical approach to Finland 1970–2007. *Journal of Transport Geography* 19: 4, 926–935.[2]

Article III Kotavaara, O., H. Antikainen, M. Marmion & J. Rusanen, (2012). Scale in the effect of accessibility on population change: GIS and a statistical approach to road, air and rail accessibility in Finland, 1990–2008. *The Geographical Journal* 178: 4, 366–382.[3]

Article IV Kotavaara, O., M. Pukkinen, H. Antikainen & J. Rusanen (2012). Role of accessibility and socio-economic variables in modelling population change at varying scale. (submitted manuscript).

[1] Reprinted from The Journal of Interdisciplinary History, XLII (2011), 89-109, with the permission of the editors of The Journal of Interdisciplinary History and The MIT Press, Cambridge, Massachusetts. © 2011 by the Massachusetts Institute of Technology and The Journal of Interdisciplinary History, Inc.

[2] Reprinted from Journal of Transport Geography, 19 / 4, Ossi Kotavaara, Harri Antikainen, Jarmo Rusanen, Population change and accessibility by road and rail networks: GIS and statistical approach to Finland 1970–2007, 926–935, Copyright (2011), with permission from Elsevier.

[3] Reprinted from The Geographical Journal, 178 / 4, 366–382, with the permission of the editors of The Geographical Journal. © 2012 Royal Geographical Society (with the Institute of British Geographers).

Contribution to papers

	Original idea	Theoretical framework	Research design	Manuscript preparation	GIS scripting and programming	Data management, database construction (and digitising)	Accessibility analysis	Statistical analysis
Paper I	OK, HA, JR	OK, HA, JR	OK, HA, JR	OK, HA, JR	HA, OK	OK, JR, HA, JH, AN	OK	OK
Paper II	OK, HA, JR	OK, HA, JR	OK, HA, JR	OK, HA, JR	HA, OK	OK, HA, JR, AN	OK	OK
Paper III	OK	OK	OK, HA, MM, JR,	OK, HA, JR	HA, OK	OK, HA, JR, AN	OK	OK, MM
Paper IV	OK	OK	OK, HA, JR	OK, MP, HA, JR	HA, OK	OK, MP, HA, JR	OK	OK, MP

Antikainen, Harri = HA

Hätälä, Johanna = JH

Kotavaara, Ossi = OK

Marmion, Mathieu = MM

Nikula, Ari = AN

Pukkinen, Mari = MP

Rusanen, Jarmo = JR

Preface and acknowledgements

The unknown is exiting. How on earth, does the world work? “Go and explore before it is too late” is an advertisement I have seen every day at work on my way to the cafeteria. In addition to travelling the World, it may be explored via computation. The project in which I began to work toward my doctoral dissertation first presented itself as simply an opportunity to do interesting work involving GIS analysis, transport aspects and spatial structure of population. After getting familiar with the analytical potential of GIS and geospatial analysis to reveal multidimensional patterns, a pinch of discovery and expedition awoke a curiosity in me towards modelability of societal issues. Exploring the world from inside the office may indeed be interesting when the adventure lies in information. So, even though the research is sometimes exhausting, it is usually exciting and the highlights of discovery are very rewarding. It is my intention that the findings of this thesis will advance perceiving the generalities in the trends in complex and multidimensional undertows of regional structural dynamics formed by the numerous decisions of individuals.

It has been my joy to work as a part of the Geoinformatics research team at the Department of Geography and University of Oulu. Our group has grown substantially since the beginning of my posting in January 2008 and the future of the team seems promising. I would like to give grateful thanks to my supervisors professor and head of our department Jarmo Rusanen, professor Miska Luoto and post-doctoral researcher Harri Antikainen. Jarmo’s pioneering work with grid-based approaches and regional structural dynamics has formed a solid basis for this thesis. I feel that Jarmo has given me a great deal of trust and freedom to take my place by recruiting me to a three-year-long project, before I even had a bachelor’s degree. Jarmo has managed our team towards a constructive and supportive atmosphere and I have always felt very welcome to visit his office with a list of questions and wishes. It has also been, indeed, good to work for a boss, who also realizes well the needs of a family. Miska has guided me to the ways of powerful statistical and geospatial analyses enabling scrutinisation of multidimensional patterns in geographic space. Non-linear statistics are an essential factor in forming the analytical basis of this thesis. Harri has supported my work with his accomplished GIS programming skills and analytical geographical thinking and also tutoring me in these fields. During our numerous analytical discussions he has answered hundreds and hundreds of questions and he has been an irreplaceable help in preparing papers.

I would like to acknowledge the work of the pre-examiners of this thesis, professor Markku Tykkyläinen and post-doctoral researcher Katarina Haugen. Their comments were very helpful, constructive and valuable for finalising the study. I have to give many thanks for several anonymous referees about their perceptive comments.

I wish to thank Mari Pukkinen for her remarkable contribution to modelling in the last manuscript and Mathieu Marmion, particularly, for tutoring me regarding the subtleties of GAMs. I wish to thank the Geoinformatics research team and particularly Heidi Määttä-Juntunen, Tiina Lankila, Tiina Huotari and Mikko Tervo for supportive and sparring discussions, for valuable comments and also being friends, not only colleagues. As far as I know, this is not a truism in the academic world. The cooperative research with Heidi (not included in this thesis) has been a fruitful experience. With Mikko, my close friend over one-and-a-half decades, I have shared much of my joys and grief, indeed.

We have had a tremendous amount of interesting conversations about geographical and particularly about non-geographical things, and many thrilling adventures too. I wish to thank, Ari Nikula and Johanna Hätälä for their substantial and precise work in digitizing historical network and population datasets in the early steps of the thesis work.

The researchers in our department have been very helpful and friendly to me. Particularly Juho Luukkonen has tutored me to navigate in the academic world and in social-theoretical issues. We have also had several interesting discussion on the nexuses of disciplines, sometimes drawing far from the surface of geographic epistemology. I remember it was Juho, who asked in 2008 if I had read the papers of Krugman. I thank Juho also for editorial work for this thesis as well as Teijo Klemettilä for layout editing. Also Petri Hottola, Pekka Kauppila, Toivo Muilu, Miia Parviainen, Juha Ridanpää, Katri Suorsa, Topiantti Äikäs and Joni Vainikka have helped much in making this thesis with discussions about general geographical topics and their specialities. I would like to acknowledge the precise and careful language check carried out by Aaron Bergdahl. Many other people have also given their valuable attention to numerous details affecting this thesis.

When beginning work on this thesis, the international cooperative network of the project felt unprecedented and challenging, but before anything promising and encouraging. Afterwards, the experience was solely good. The major work for this thesis was carried out during the project FiRa (Finnish Railways in the Nordic and Russian context), which was funded by the Academy of Finland during 2008–2010. FiRa was a cooperative sub-project of ESF's (European Science Foundation) project WRR (The Development of European Waterways, Road and Rail Infrastructures: A Geographical Information System for the History of European Integration 1825–2005). I wish to thank WRR participants and particularly the project leader, professor Jordi Martí-Henneberg, for fruitful discussions from different European perspectives and different methodological bases.

Another strong international network that greatly advanced this thesis is the ESPON TRACC (TRansport ACCessibility at regional/local scale and patterns in Europe) project. I wish to thank TRACC participants for their very professional cooperative work, and especially the project leader, Klaus Spiekermann, for several essential remarks considering accessibility analytical issues and regional structural dynamics improving this thesis remarkably.

Some aspects of this thesis were applied to the new grid-based regional typology of Finland. I wish to thank former director of regional development Janne Antikainen for striving for this typology and also promoting the inclusion of accessibility aspects to it.

Finalising this thesis has been possible by funding from the Department of Geography at the University of Oulu, the Geography Graduate School, the Foundation of Science at Oulu University Scholarship Foundation, the Foundation of Tauno Tönning and the Foundation of Emil Aaltonen, which all are wished to be acknowledged.

I wish to thank my friends, parents and most of all, my wife. Life is so much nicer when it is explored together with friends, and they have really shared much of my feelings during the work for this thesis. My parents have continually encouraged me in my path to find an interesting education and they have supported me greatly during my studies. They have also been very helpful during the last year, after the birth of my son. Finally, Niina, my beloved wife, you have tolerated the highlights, worst moments and the thousands of everyday issues of preparing this thesis. Thanks cannot be enough for this. Also, with a great regularity it has been my work, not yours, which has got the emphasis – I hope that I can give the same support to your thesis. I feel that without your reliable support, I would not be at this point, neither in my life nor my studies. I would feel silly dedicating this thesis to you and our lovely son Joel, since perhaps you would prefer that I dedicate my time.

(In the figurative space of the final stretch)
At my office at the University of Oulu
1. November 2012

Ossi Kotavaara

1. Introduction

The aim of this thesis is to explore the effects of transport accessibility on population change at different scales in Finland between 1880−2009. The analytical framework of the study is founded on geographic information systems (GIS) as technology and also on the geographical information science (GIScience) behind it (Brown *et al.* 2004; Goodchild 2004). According to Longley *et al.* (2005: xi) the field of GIS is concerned with the description, explanation, and prediction of patterns and processes at geographical scales. The geospatial analytical ensemble of the study consists of accessibility analysis implemented on GIS (Miller and Shaw 2001), which were also applied to data management, and of non-linear multiple regression, generalized additive models. Theoretical interests rise from the interface, or nexuses, of transport and economic geography and regional science. Exploring the potential relationships between accessibility and population change at different scales is possible due to the availability of GIS data of Finnish transport networks and the population having high spatial accuracy and also a long historical reach.

This study adopts an empirical-statistical approach to modelling the investigated phenomena. As a distinction from analytical-theoretical and mechanistic-process models, this approach does not attempt to describe realistic causalities or inform about underlying functions and mechanisms; instead its objective is to explore and condense empirical facts (see Guisan and Zimmermann 2000). Indeed, the attitude towards studied phenomena and interpretation of the results in this thesis have associations with the thoughts of Fotheringham *et al.* (2000: 5):

> *"In (quantitative) human geography, where the subject matter is typically clouded by human idiosyncrasies, measurement problems and uncertainty, the search is not generally for hard evidence that global 'laws' of human behaviour exist. Rather, the emphasis of quantitative analysis in human geography is to accrue sufficient evidence which makes the adoption of a particular line of thought compelling."*

The scientific motivation towards this study originates in the empirical-analytical and theoretical gaps in the field of accessibility studies. The reasoning behind the relationship between accessibility and population concentration makes intuitive sense also in Finland. On the basis of an extensive and recent review of transport studies (see Spiekermann *et al.* 2011), there is an apparent lack of research combining the measured accessibility with population change using statistics, particularly in Finland. Also the effect of accessibility is usually addressed separately for different transport modes, instead of taking them into account simultaneously. The explorative perspective of the study is mainly basic research, but the results also have applied characteristics. At the centre of the study are the development trajectories in the regional structure and the history of transport and population in Finland; exploring these with accurate and high-quality transport, population and socioeconomic datasets; and taking new, more accurate insights by non-linear regression analysis, generalized additive models (GAMs).

In numerous studies, accessibility and population change have been analysed at various different spatial levels as separate phenomena or by examining their connection. However, there is still a lack of knowledge about the relationship between accessibility and population change and particularly about how it is scale dependent. Watson (1978) has observed that geographers tend to work at one analytical level, exclusively and implicitly, without considering other alternatives. Changes in scale change the importance and relevance of variables (Meentemeyer 1989: 165) and the choices over scale, extent, and resolution in the analysis may critically affect the type of patterns that will be observed (Gibson *et al.* 2000: 221). Also, problems related to ecological fallacy and spatial autocorrelation (see de Smith *et al.* 2007, 62–63, 95–99) may be encountered in relation to the scales of the studies. Verburg *et al.* (1999: 47) concludes well that coarse scales are usually useful in revealing general trends, but due to a high level of aggregation, these coarse scales can obscure the variability, and they are, thus, inaccurate for fine scale and local assessments. Population change is, indeed, an outcome of numerous individual location choices and the scale on which these choices may form a pattern in the context of transport accessibility is unclear. Hence, this thesis takes the claim of Marceau (1999) into consideration, one that it is necessary to identify scale thresholds to understand the interactions that are occurring within and between the levels of organisation. In addition to considering the effect of accessibility at different scales, also key socio-economic characteristics are included in the analysis.

The struggle to increase accessibility is continuous, while different actors (e.g. individuals, enterprises, regions or states) would like to improve their ability to transport or to be reached. Thus, political and policy relevance underpin the legitimacy of the theme considered herein. The importance of transport within the context of political agendas originates from the effect of a transport system on productivity, trade and location choices, contributing in turn to the varied growth of regions (Rietveld and Bruinsma 1998). Since the early days of Finnish transport politics, there have been highly differing aims and motivations for investing in transport systems and the regional and modal competition has indeed been hard (Seppinen 1992). Presently, for example, the Finnish government published the political aims of the national transportation policy (Kilpailukykyä ja hyvinvointia... 2012). The transport policy and the transport system are seen to be tightly connected to the other functions of society, including, in particular, the requirements of industry, the economy and employment, as well as regional development, since good accessibility is a key factor in the economic development and prosperity of regions. In contrast, improving the accessibility of remote rural areas will not be the focus of the Finnish transport policy. The planning of the transport infrastructure has been deeply integrated into Finland's planning system, ranging from the municipal and regional levels, to the scale of the state (MRA 1999; MRL 1999; Valtioneuvoston päätös... 2000; Tarkistetut valtakunnalliset... 2009). Also in the European Union, policies promoting transport accessibility have an essential role in developing regions (European Commission 2011).

Finland is an interesting case for research, considering population change in the context of transport accessibility, for at least five reasons. First, from a historical perspective, the Finnish population was almost evenly distributed before the inception of a large-scale transport infrastructure, and the remarkable urbanisation process that simultaneously began with motorisation (Alestalo 1983). The case of Finland is an opportunity for the understanding of urbanisation in sparsely populated and peripheral countries, which differs from the typical setting of urbanisation in Western Europe and America. Second, because of the remote northern location and the border with the Baltic Sea, Finland has had limited connections with other countries, in comparison with Europe (see Spiekermann, and Aalbu 2004). Therefore, in spite of the needs to concentrate activities in the southern regions, which have a better reach to foreign and domestic markets, the effect of international connections on the Finnish regional structure may be considered to be rather weak. Third, the regional and urban structures, characterized by sparse population, small cities and towns, and long distances, differs greatly from the rest of Europe (*idem.*). Fourth, the concentration trend within the scale economies is evident also in Finland (Lehtonen and Tykkyläinen 2012). Finally, a substantial benefit to the research is gained from availability of excellent quality population and transport network data enabling accurate scale and long-term analysis.

Even though significant efforts have been directed at accessibility research in general, the transport geographical and accessibility analytical research tradition in Finland is in its early stages. Only a few scientific studies have focused on accessibility in Finland. Vuoristo (1967) has assessed the Finnish road network connectivity with quantitative network analyses. Hakala (1973) has computed accessibility indices on the basis of cost distances to define service accessibility in the context of central-place theory. Tykkyläinen (1981) has analysed the structure of Finnish provinces by applying a set of accessibility indices. A deep historical perspective of the development of highway capital 1900–2009 is given by Uimonen (2010) with an accessibility analysis applying to municipalities as destinations and 10×10 km grid cells as origins. Helminen *et al.* (2012) have analysed urban-rural interaction using the accessibility approach, defining distance decay patterns in Euclidian space. Toivonen *et al.* (2010) have assessed multimodal accessibility at the grid cell level in the Helsinki region by including multimodal bus, train tram, metro and walking in calculations. Finland is included in the grid cell-based study of Schmitt *et al.* (2008), which analyses the Baltic Sea Region accessibility potentials and the accessibility of selected transport and education facilities. The accessibility of Finland as a Nordic periphery is considered in the study of Spiekermann and Aalbu (2004), where road, air and rail accessibility is assessed at the municipal level. Further, Gløersen (2009: 46), assessed the travel-times to towns with more than 10 000 inhabitants in peripheral Finland, Sweden and Norway.

It is very common that the effects of other factors are ignored in studies considering the effects of accessibility, probably due to the limited availability of proper variables. Thus, an omitted-variable bias may occur widely. As geographic space is usually ignored in demographic research, there is a need to explicitly bring spatial processes into empirical demographic studies to correct any potential misspecifications (Voss 2007: 468, 471). Traditionally, population change is analysed with birth and death rates connected to net migration and possible correction factors. The research in this field has produced an ample set of highly sophisticated statistical models (see Alho and Spencer 2005; Booth 2006). In contrast, land-use dynamics, such as urban growth, are modelled by spatial variables, also including distance-based indicators (Andersson *et al.* 2006; Liu and Seto 2008; Pijakowski *et al.* 2002; Verburg *et al.* 2004). Scrutinising spatial characteristics of population change in the context of accessibility may thus give information usable in improving spatial and temporal accuracy of population change models.

A number of demographical studies have considered population trends in Finland, and extensive reviews are given by Pitkänen (1988), Aro (2007) and Koskinen *et al.* (2007). An early example of geographical analyses focusing on Finnish population change is the study of Alestalo (1983), considering concentration at the municipal level with a deep historical perspective. Population change is analysed with accurate grid cells and GIS by Halme (1999) and Rusanen *et al.* (2003), with aims to explore spatial dynamics of population and to define the changes in the regional structure. Kauppinen (2000) has applied grid-based approaches to the study of migration. Migration process is studied with descriptive statistics and cartographical analyses by Heikkilä (2003) with emphasis on long-term process and by (Heikkilä and Pikkarainen 2010) focusing on concentration and related policies. Despite dedicated research efforts aimed at understanding population dynamics, spatial models considering population change in Finland are commonly demographic trend calculations at the municipal level (Population projection 2012).

2. Accessibility and its effects on population change

Transportation is generally thought of as the way to reach or move something through space (Black 2003: 3) and accessibility refers to the ability to transport. Some exact definitions are given regarding the analytical use of the concept of accessibility. Spiekermann and Wegener (2007) note that accessibility is the main product of a transport system. According to Chen *et al.* (2007), accessibility typically refers to the 'ease' of reaching opportunities for activities and services and can be used to assess the performance of a transportation and urban system. Holl (2007: 286) states that accessibility expresses the extent to which spatial separation can be overcome. Geertman and Ritsema van Eck (1995) define accessibility to be the ability to command the transportation facilities that are necessary to reach desired locations at suitable times. Understanding accessibility in this thesis has its most solid basis in the definition given by Geurs & Ritseman van Eck (2001: 19):

> *"Accessibility… …the extent to which the land-use transport system enables (groups of) individuals or goods to reach activities or destinations by means of a (combination of) transport mode(s)."*

Concentration has been the major trend in the change of population distribution across the world. Cities are seen to exist because they offer opportunities for comparative advantage increasing trade, the economies of scale resulting internal firm savings and the economies of agglomeration leading to external location advantage (Malpezzi 2011: 49); because they foster high levels of access (Taylor 2004: 299) and to eliminate transport costs for people, goods and ideas (Glaeser 1998: 140). In general, the dynamics of population distribution are affected largely by the economy; according to Glaeser (1994: 18), most things that predict per capita income growth also predict population growth.

Since the formulation of early theories of urban spatial structure (*e.g.* von Thünen 1826; Alonso 1964) and industrial location (*e.g.* Weber 1909) in the context of distance, accessibility and its effects have received particular attention in numerous studies in the field of transport geography, but also within economic geography, as the space has striven to be included in economic models during the last couple decades. Various analytical approaches and different types of indicators are applied to a number of case study areas, and extensive reviews have been produced on the subject (Geurs and Ritsema van Eck 2001; Rietveld and Bruinsma 1998; Rietveld and Nijkamp 1992; Spiekermann *et al.* 2011; Wegener and Fürst 2004).

There has been no full agreement on the effects of accessibility, but some general principles or key findings may be promoted. The agglomeration of activities may be explained by the economies of scale, positive feedback in circular causation and increasing returns (see Fujita *et al.* 1999). According to Krugman (1991), emergence of a core-periphery pattern depends on transportation costs, economies of scale, and the share of manufacturing in national income. Fujita and Thisse (1996: 368) state even more specifically, that there is a fundamental trade-off between scale economies and transportation costs in the geographical organisation of markets and low transport costs tend to favour the formation of geographical clusters or to deter the creation of new ones.

The importance of accessibility to economic development and urbanisation on a local and regional level is evident (Biehl 1991; Graham 2007; MacKinnon *et al.* 2008; Quinet and Vickerman 2004: 1–70). Accessibility can be considered to have a contribution to the concentration of economic activities and the population (*e.g.* Herzog and Bjornstad 1982; Rietveld and Nijkamp 1993; Wegener and Bökemann 1998; Quinet and Vickerman 2004: 1–70). This is, however, possible only if the transport system is efficient enough and the low level of transport costs enable the concentration and emergence of the core-periphery pattern, together with increasing returns and changing demand (Krugman 1991). In accordance with this theory, in a pre-railroad, pre-motorized and preindustrial society, the population is limited, only a small fraction of the population works in manufacturing, and high transportation costs ensure weak economies of scale. In societies with mass consumption and production, in which transport costs are low, the economies of scale are evident and the concentration may begin along with the location and re-location of the production and people. According to Vickerman *et al.* (1999), the relative gains in accessibility of peripheral regions may be beneficial to their economic development, but these gains will always be over-shadowed by the much larger gains in accessibility of the regions in the core. In addition, the concentration pertains to transport networks too, as European policies tend to favour transport investments improving the accessibility of urban centres (*idem.*), and the trend has also been evident in Finnish transport investments in Finland (Transport policy guidelines... 2008).

The role of accessibility is rather clear in the general development in the centre-periphery pattern, but consumption amenities may greatly affect the success of competing for a population (Rappaport, 2008; 2009) In addition, a remarkable variation exists at an individual level, where attention should be paid to personal needs and wishes instead of 'objective' measurements, as a certain residential characteristic may be considered to be both an advantage and a disadvantage by different people (Haugen 2011). While the effect of accessibility on a regional level depicts the concentration of the population to centres, accessibility also has dispersing effects on smaller scales. The pattern of the concentration may be stirred by the trends related to counter urbanisation, in which Mitchell (2004) identifies the differing trends of anti-urbanisation, displaced urbanisation and ex-urbanisation. In terms of accessibility, the question pertains to the two latter phenomena, as accessibility clearly affects the population growth on the urban fringe (Brueckner 2000; Nechyba and Walsh 2004). This phenomenon, called urban sprawl, is an essential process that influences the population distribution virtually everywhere in the

Western world, especially in the United States. The trend is related particularly to increasing motorisation, under conditions of increasing incomes and employment (Glaeser and Kahn 2004; Patacchini and Zenou 2009). In sparsely populated and recently urbanized Finland, the sprawl trends started to intensify during the 1990s. Vartiainen (1989) has considered sprawl, or regionalisation, trends in Finland, mainly not as a problem, but rather as a potential development trend towards competitive large urban regions.

The importance of accessibility to regional development is an indisputable fact, however, the actual causality is complicated and indirect (see Wegener 2004). In addition, the effects of accessibility are not completely straightforward (Spiekermann and Wegener 2006) and the study of Banister and Brechman (2001) explicitly shows that the growth process within accessibility is only possible with the support of political, policy and institutional factors. They also see that the benefits of increased accessibility that exist in one location occur (potentially) at the expense of a competing location.

Several studies underpin the agglomeration trend in society, by relating accessibility to dynamics in the centre-periphery pattern (Johansson *et al.* 2002; Song 1996; Spiekermann and Neubauer 2002). Accessibility is also in a key role in the location choices of industrial parties (de Bok and van Oort 2011; Klaesson and Johansson 2008; Song *et al.* 2012). In addition to the attractive effect of accessibility which is commonly present in centres, densely populated industrialized areas tend to have a higher network infrastructure endowment in comparison to peripheries (Nijkamp 1986, 15). As a curiosity, railway accessibility has been observed, in some cases, to decrease the attractiveness of rural areas to migrants (Vaturi *et al.* 2011). However, the effects of accessibility may vary interestingly across space. For example, in the study of Chi (2010a) the accessibility of the highway network had an increasing effect on suburban populations, but there was not any effect to in centre areas. Again, in another study by Chi (2010b), airports were noticed to benefit the growth of suburbs, but not of centres. Hence, studies concerning the relationship between accessibility and population change in Finland are important.

3. Research design

The aim of this thesis is to analyse the effect of transport accessibility on population change in Finland on different spatial scales. Almost the entire period of modern transport in Finland, covering the years 1880–2009, was investigated during the study, by laying focus on the sub-periods 1880–1970, 1970–2007, 1990–2008 and 2003–2009. The study made use of high-quality digital transport infrastructure models and accurate register-based population grid data (Grid database 2004; 2010) that are free of administrative boundaries. The historical analysis of population change in this study required the creation of GIS databases concerning the population at the municipal level and the railway network. Road accessibility was analysed by gravity idea-based accessibility potentials, which were also tested for multimodal road and rail accessibility. The accessibility of railway stations and airports were assessed as the travel time from population location nodes. Accessibility computations, spatial analyses and data management were carried out by using geographic information systems (GIS). Accessibility variables were related to population change with non-linear regression, generalized additive models (GAMs), which also formed a basis for considering the importance of accessibility variables. Eight scales of analysis were used, beginning with municipal and built-up area scales, continuing to grid cells at six resolutions. The 2×2 km grid cells was the most accurate resolution at which the calculation of accessibility potentials was computationally possible and the largest 24×24 km grids are roughly comparable to the municipal level in terms of resolution. In addition to taking examples to evaluate the consistency of models in space, temporal consistency was also considered. Finally, the relevance of accessibility variables in modelling population change was assessed within a careful selection of socio-economic variables, which were included in the models.

The research questions in this study are:

1. How does potential accessibility explain population change occurring in the context of a regional and urban structure?
2. What are the most applicable distance decay functions for accessibility potentials to explain population change at different scales?
3. What is the effect of scale (*i.e.* size of spatial units or resolution) on the relationship between accessibility and population change?
4. What are the effects of a) rail and air transport infrastructure accessibilities, b) key socio-economic variables c) and population density to population change at different scales?
5. Is it possible to predict population change spatio-temporally, by extrapolating the model results?

The analytical work is performed in research papers and the correspondence to the research questions is presented in table 1. In paper I, examinations were established for road and rail accessibility on a municipal scale (1880–2007) and also on a built-up area scale in paper II (1970−2007). In paper III and IV, a six grid cell resolution-based analysis was established to cover the period 1990–2009 and the accessibility of airports is also included in the analysis. In paper III, the temporal consistency was tested by predicting the 1990–2008 population change with a model calibrated for the period 1990–2000. The models covered population change as relative and absolute numbers. In paper IV, key socio-economic variables, selected from a group of 58 variables, are tested in a model covering the period 2003–2009. According to the literature survey, the effect of the accessibility in relation to population change studies has not been previously investigated with such an extensive and systematic study setting.

Table 1. Specified contribution of papers in relation to research questions.

	Paper I	**Paper II**	**Paper III**	**Paper IV**
Time period	1880–1970	1970–2007	1990–2008	2003–2009
Question 1	From the growth of incipient cities to rapid agglomeration to accessible urban areas in the context of major structural change.	Differentiating periods of conservative municipal scale urbanisation and intensifying agglomeration within accessible municipalities and built-up areas	Differentiating regional growth and urban sprawl	Modelling population change dynamics in the context of accessibility and key socio-economic variables
Question 2	Linear and quadratic distance decay functions	Linear and quadratic decay functions	Linear and quadratic decay functions	Linear, quadratic and four negative exponential, including empirically estimated decay functions
Question 3	Comparisons of potential accessibility and population density on a municipal scale	Municipal and builtup area scale	Six grid cell resolution between 2×2 km and 24×24 km.	Six grid cell resolution between 2×2 km and 24×24 km
Question 4	Accessibility of rail transport facilities and population density	Accessibility of rail transport facilities	Accessibility of rail and air transport facilities and population density	Accessibility of rail and air transport facilities, socio-economic variables and population density
Question 5			Full dataset and small constant size random sample-based model validation	Full dataset model validation and four times cross-validation

4. Key geographies and history of Finland as the study area

This chapter considers the development of transportation and the development of population and socio-economic trends in Finland. During the research period, Finland has changed from an agrarian grand duchy to a post-industrialized economy being a member state of the European Union. The land transport system has develop a great deal from the early stages, when manually powered local transports were enhanced with steam powered railways creating enormous increases in transport capacity within the reach of the network. The emergence of combustion motor-based travel and transport led to motorisation. After World War II and particularly during the latest decades of the study, air transport has become a very remarkable transport mode.

4.1 Transport system

Finnish roads have a long history as a ubiquitous network of paths, cartways which served in historical transports in conjunction with waterways. By the beginning of the nineteenth century, the road network covered most of the main towns of the country (Suomen teiden historia 1974). Lakes, rivers and later canals were also important for travel and particularly for larger transports (Seppinen 1992: 9−12), but were constrained largely by topography. Hence, the transport system improved a great deal when mechanically powered railways were established. Zetterberg (2011) presents an extensive history of Finnish railways, in which railway line constructions were, indeed, impelled by numerous local interests when general motivations varied from improving connections to develop trade, industry, education and forestry policies to design a transport system to be applicable for first Russian and then Finnish security policies. However, it can be concluded that the most important motivation was to connect the principal population centres, but, in addition, many smaller places gained increased accessibility by railways (Valtionrautatiet... 1987 [notice the wrong time period in the title of the reference]; Seppinen 1992).

The first railway link in Finland was established in 1862, the railway connection to Russia was opened in 1870 and before the turn of the century, three north–south-directed trunk lines were opened to traffic. Interestingly, when small and large towns were connected by railway, the benefits were larger in the latter than the former (Zetterberg 2011: 86). The first cars appeared on Finnish roads before the turn of the century, busses a bit later and, finally in 1918, the state took over road maintenance from farmers (Seppinen 1992: 82−83). The railway connection to Sweden opened in 1919. However, the Finnish rail gauge complies with the wider Russian variety, which differs from the European standard used in Sweden. Thus, the connection between these rail networks has not ever been fully functional. By being the only serviceable motorized land transport mode during the nineteenth century, railways achieved dominance, which continued until the mid-twentieth century. Due to limited local coverage of the railway network, roads certainly maintained their role as connectors at the local level. The significance of the road network began to rise again during the 1930s, when more flexible coach traffic began to seriously compete with railways. During the early twentieth century, the state extended the trunk-line network almost to its present reach. Some privately owned and operated lines were established for local traffic, extending their outreach and introducing regional branches, but the success of these companies was commonly very limited (Valtionrautatiet... 1987).

The emphasis of motorized passenger travel shifted from railways to busses between the years 1934–1939 (Yearbook of Finnish… 1958) and to passenger cars from 1957 (Yearbook of Finnish… 1960) measured by the passenger kilometres travelled. The motorisation of passenger vehicles began to gain momentum during the 1950s. The main Finnish road network was developed simultaneously with the growing number of cars to satisfy increasing transport needs and the major peak in road investments took place during the 1960s, when motorisation escalated in full. Speed limits were initially determined during the 1920s, but in practice between 1938−1973, the driving speed was unlimited in rural areas (Seppinen 1992: 88−90, 167–172). At the beginning of the 1930s, Finland had about 20 000 automobiles on the road, in 1950 the number was closer to 27 000 and in 1960 the number had increased to around 183 000 (Yearbook of Transport… 1980: 58). The increase to 711 000 cars by the year 1970 was huge, whereas the number of cars in 2009 was 3.2 million (Transport and communications… 2010: 26). Passenger cars have resulted in more person kilometres travelled than busses since 1960 (Yearbook of transport… 1971). Along with motorisation, interest in expanding the railway network decreased. The primary objective for developing railway transports shifted to facilitate long-haul connections and local traffic within the capital-city area, although some new lines were still opened with motivations originating from regional subsidizing. By 1970, the Finnish railway network had almost achieved its saturated reach. The only remarkable exception is the short-cut link, which was established in 2006, reducing travel times between Helsinki and Lahti (Figure 1).

Bus transports experienced their most intensive period during the 1980s in passenger kilometres (Yearbook of transport... 1990: 18), after which their absolute and relative share of travel has diminished strongly due to passenger cars. (Yearbook of Transport… 1980: 5). Bus transport has been mainly effective in urban areas, due to sparse population densities in peripheral areas. The share of passenger transport by railways has been relatively limited during the research period of this study. During the 1970s, the railways extensively served local traffic, but decreasing effectiveness and increasing motorisation forced railway transports to be rationalized by improving long-haul connections and by focusing on commuter traffic in the capital city region (Valtionrautatiet... 1987: 37–51). Long-haul transports were improved after 1995 in particular, due to investments in faster trains operating in intercity connections. The counter processes for this have been the

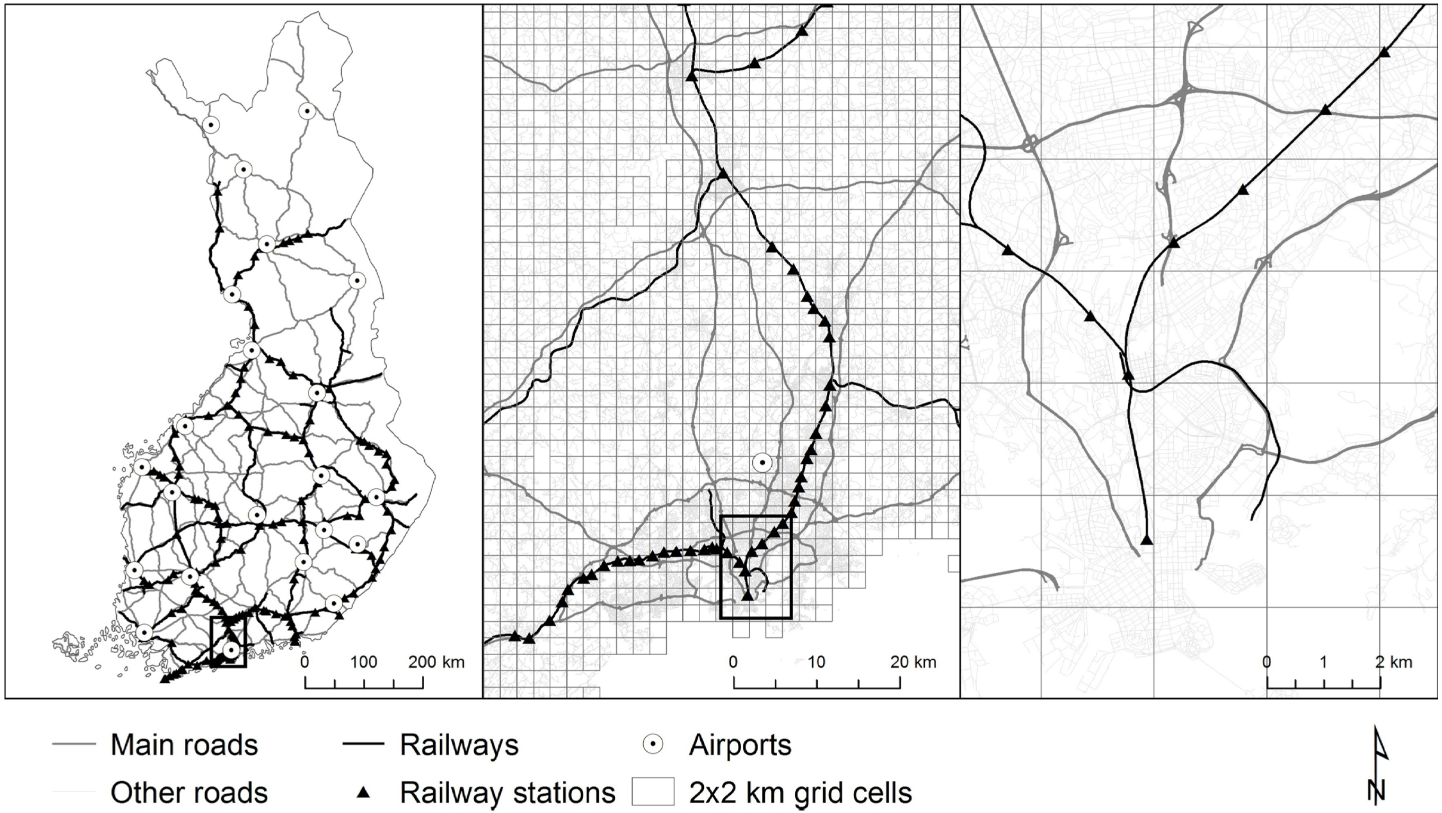

Figure 1. Different scale examples of spatial structure of Finnish transport networks in 2003 and accuracy of data.

huge reduction in the number of peripheral railway connections (Valtionrautatiet... 1987: 29–30). Within four decades the number of railway stations has decreased to one-tenth of the original number (Finnish railway statistics 2010: 25). Railway transport has been dominated by a state-owned company, unprofitable in pure economic terms, between the years 1955 and 1990 (Railway statistics 1989, 1990). Therefore, prior to the 1990s, railway transport had effectively been a form of regional subsidy rather than an actual business activity.

Air transport began to be an actual mode of transport by 1937 and during the 1950s, the air transport network could be characterized as being extensive and active (Seppinen 1992: 102, 155). Over the past two decades, civil airports have covered the centres of Finland well and they have also greatly improved accessibility in some deep peripheries. Since the year 1990, over 20 civil airports have been in active operation. Helsinki Airport has been the air transport hub of Finland for decades. It was used by 73.6 % of Finnish air transport passengers in 2003 and 78.0% of Finnish air passengers in 2009 (CAA's Air traffic statistics 2004; Finavia's Air traffic statistics 2009). Helsinki Airport is the only airport performing very well and its result partly funds a majority of Finnish civil airports (Finavia Oyj Vuosikertomus 2011: 50), which is a form of a remarkable subsidization.

4.2 Population and socio-economic trends

Since economic, investment and political conditions are essential factors, enabling the development of the economy related to improving the transport system (Banister and Berechman 2001), the main trends of the economy and regional development are illustrated here. In European perspective, Finland can be characterized as a northern and peripherally situated country behind the Baltic Sea. Finland borders Russia to the east, Norway to the north and Sweden to the north-west. During the research period, the population of Finland increased from 2.1 million in 1880 to 5.4 million in 2009 and the total population density increased from 6.8 per km² to 17.6, respectively (Population by industry 1979; Grid database 2010). In inhabited grid cells, the population density has increased from 41.3 per km² in 1970 to 52.0 in 2009 (Grid database 1970; 2010). Finland has some archipelagic areas, such as the Åland Islands, which are excluded from this study due to their marginal share of the population and their highly differing transport system, which is based on ferry connections.

In the early nineteenth century, Finland could be characterized as being almost completely agrarian, immobile, and staid, and migration took place mainly at a local and sub-regional level (Aro 2007: 88–94). Industrialisation began during the latter half of the century (Rasila 1982: 132) locating mainly in towns and rural centres (Aro 2007: 93) and migration to cities intensified (Alestalo 1983; Pitkänen 2007: 65). Prior to 1880, migration was restricted from people in non-possessing standings in society (Aro 2007: 268). In 1880, the primary source of livelihood was agriculture at 74.9 % and the share of industry was only 5.3 % (Population by industry 1979). Around the turn of the century, in rural areas, there was a remarkable out-migration pressure and industrialising towns had a deficiency of labour, which increased migration flows towards regional centres (Aro 2007: 92–93). Political acts were implemented to form small estates from crofts and state lands. However, around the turn of the century and especially during the 1910s, a number of inhabitants, amounting to about 415 000 people, also migrated overseas, mainly to the United States and Canada (Vahtola 2003: 300–302). The 1920s were characterized by strong economic growth and industrialisation, which fell in the recession that reached Finland around 1930 (Ahvenainen and Kuusenterä 1982: 223). The following major population trend occurred during and after World War II. More than 400 000 people from areas that were seized by the Soviet Union re-settled in a wide range of municipalities (Vahtola 2003: 383).

The urbanisation process began to slowly develop in the last decades of 19th century (Koskinen *et al.* 2007) During the 1960s, immigration toward cities accelerated to a level which had never been reached before in Finland (Pitkänen 2007: 65). Simultaneously, particular policies were established to balance regional developments (Aro 2007: 150−151). A remarkable number of migrants also moved to Sweden and out-migration was so intensive, that population in Finland decreased during 1969–1970 (Pitkänen 2007: 63). In comparison to other Western countries in general, the Finnish population has experienced a relatively late urbanisation, beginning to a significant extent after the post-war era and peaking during the 1960s. In 1970, the population was mainly employed by commerce, transportation, and services (32.3 %), as well as by the industrial sector (22.8%). The share of workers in agriculture had been decreased to 17.6% and the share of construction was 8.7 % (Population by industry 1979). Since the 1970s, the urbanisation process was connected to the shift from an agricultural society to an industry-based economy, leading to the depopulation of peripheral areas.

Until the early 1990s, the economy of Finland grew substantially, which enabled large growth in the public sector, and regional competition was balanced by remarkable subsidizations to undeveloped regions. Since the beginning of the 1980s, population growth pressure shifted towards margins of urban areas and to areas that were adjacent to centres, at first, mainly emphasising the capital area, but then also regional centres (Heikkilä 2003; Rusanen *et al.* 2003). The population in the peripheries concentrated on a local level towards town centres, but the surrounding areas degenerated rapidly during both the 1970s and 1980s (Halme 1999: 54–71). During the 1990s, several structural changes occurred in Finland, including a deep recession caused by the collapse of trade with the Soviet Union and European Union membership in 1995, which improved access to a larger European market, as well as a transition towards information and communication technologies (OECD Territorial Reviews... 2005). Regional subsidies to the peripheries decreased and efforts shifted to developing centres (see Tervo 2005). Since the 1990s, the strongest migration flows, characterized by sprawl-like centripetal trends, have mainly been aimed towards the southern parts of the country and, particularly, to the largest cities and their environments. Simultaneously, areas located remotely in relation to growth centres have experienced declining populations (Rusanen *et al.* 2003). By 2009, Finland had clearly become an urbanized country (Figure 2) and the post-industrial economy, portrayed by 74.1 % population being employed by the public sector and services. Industry, however, had an important role (14.7 %) in comparison with construction (6.3 %) and primary production (3.7 %) (Työlliset toimialan... 2012).

The population dynamics have naturally also been influenced by the politics, systems of governing and planning that are characteristic of each historical period. These have, indeed, varied a great deal during the research period, just like the whole society. However, the extensive presentation of this composition and its development is unreasonable within the frames of this thesis, due to the remarkable complexity and needed profoundness. However, the basic structure of the present system allocates a considerable amount of power to the local level in municipalities, which to some degree are self-governing units in the field of planning (MRA 1999; MRL 1999; Valtioneuvoston päätös... 2000; Tarkistetut valtakunnalliset... 2009). It is clear that with planning, local population patterns within neighbourhoods can be effectively controlled, but their effect on population flows between municipalities is only indirect.

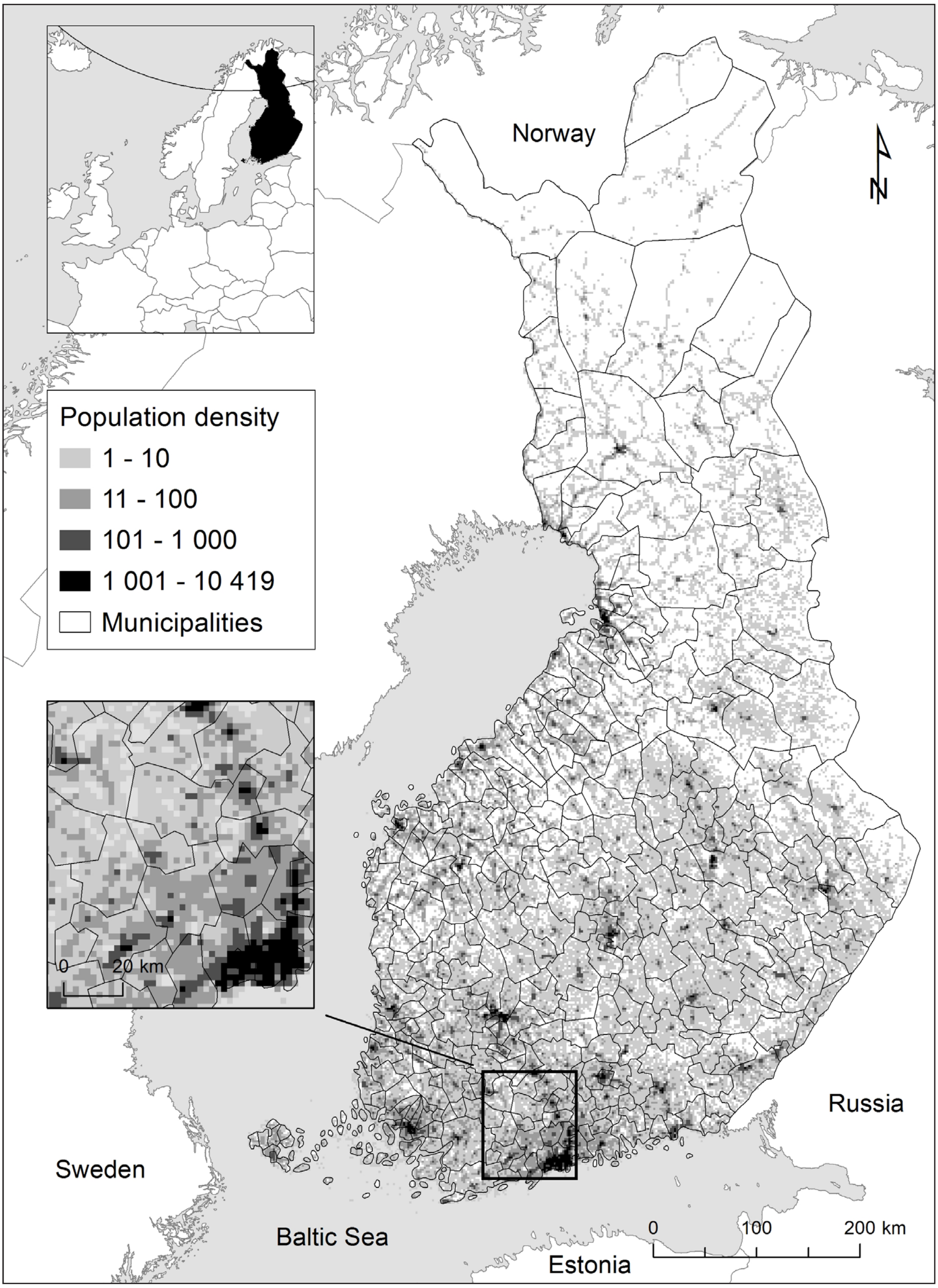

Figure 2. Finnish municipal structure in 2007 and population distribution in 2008 (inhabitants per square kilometre at 2×2 km resolution).

5. Data and methods

5.1 Data

The data of this thesis consists of transport, population and socio-economic datasets. Transport data covers roads, railways with stations and airports. The road network has an accurate geometry and attributes for 2003 and for previous times different types of estimates are used. Rail network data has good temporal quality, as several attributes and also station data are attached. Eight scales are used for population datasets. Municipal-level data is available for the entire study period, whereas population grid data is only available since 1970.

5.1.1 Transport network and facility datasets

The GIS data of the road network was obtained from the Digiroad database involving speed limit information for commonly used roads (Digiroad 2012). The earliest available version of Digiroad represents the 2003 road network. It was used in computing the travel time estimates in papers III and IV. To estimate travel times accurately enough, all the regularly used roads were included in the analysis. The used road network data includes regional and local main streets, collector and feeder streets and private roads allowed for public use. The database has no historical attributes. Hence, for paper I, the higher road classes were selected and travel time attributes associated with the road links were modified to reflect the assumed travel conditions during different eras, providing a rough estimate of the historical accessibility. The approach has uncertainties in accuracy, but remarkable benefits in comparison to direct distance estimates ignoring, *e.g.,* water bodies. For paper II, higher road class geometry and speed limits were digitized on the base of historical maps covering all main and regional roads from 1973 onwards. The state-scale maps were provided by the Finnish Road Administration.

In Finland, congestion is not generally regarded as being a factor affecting the daily accessibility between municipalities or built-up areas, with the exception of some occurrences within the limits of the capital city area (see Kalliokoski 2003). Therefore, congestion effects are not incorporated in the data more specifically than that which is already expressed in speed limits in paper II, the analysis of which is carried out at municipal and built-up scales. To estimate the effects of congestion for the grid cell-based analysis of paper III, travelling speeds were corrected by a factor of 0.8 in built-up areas. The geometry and capacity of the main and regional roads was relatively stable during the research period, as emphasis was on maintenance. This was apparent in the stability of the average vehicle speeds (Uimonen 2010).

Inland islands were included in the analysis, and as there are no significant waterway connections in continental Finland, short road ferry connections, which are included in the digital road network data, were the only water transport routes that were included in the analysis. However, because of their highly differing infrastructure, relying on ferry connections, the coastal archipelagos and the Åland islands had to be excluded from the analysis.

The GIS data of the railway network was constructed to represent spatial and temporal development. The railway network geometry for 2007 was provided by the Finnish Rail Administration. The attachment of relevant attributes, primarily the opening dates of links, estimated travel times, and rail classifications and also locations of stations and stops, was carried out by means of linear referencing (see *e.g.* Verbyla 2002) and digitising. The use of linear referencing was possible due to the high level of the temporal permanency of the network and the railway kilometre-based positioning system used by the Finnish railways from the outset. The attached attributes were compiled from literal documents. For example, the kilometre positions of railway stations and the use of stops at different times was determined from engine-driver timetables and the usage of stations from the passenger timetables. Narrow-gauge railways are excluded from the study because of their local role and their limited connectivity to the main railway network. The stations were connected to the road network by automated digitising. The GIS data for airports was compiled from AIP Suomi – Finland (2012).

5.1.2 Population and socio-economic data

The representing population-change variable was treated as a relative number. Because the population-change values did not follow normal distribution when measured in percentages, the logarithmic function of the relative population change had to be used instead to gain better correspondence. The equation for population change can be written

$$\Delta P_{\log} = \log \frac{P_t}{P_{t-\Delta t}} \qquad (1)$$

where ΔP_{log} is the relative logarithmic population change, P_t and $P_{t-\Delta t}$ are the populations at the end and the beginning of a period, respectively.

In this thesis, population change is assessed at a municipal level in papers I and II. Municipal data is used due to its extensive temporal reach beginning at the year 1880 with decadal intervals and its role as a basic spatial division for assessing the population and other trends in Finland. The municipal structure of Finland has been subject to change throughout the research period. Statistics Finland has produced temporally comparable municipal population data since 1951. The data which was applied in paper I and II was available in the 2007 municipal boundaries. The population figures covering the period 1880−1940 had to undergo manipulation to ensure temporal comparability over the decades. The database was constructed to follow the municipal structure in 2007. The population numbers were compiled from decadal tables which also included descriptive notes about territorial changes in the municipal structure affecting populations. As that accurate information regarding municipal boundaries in the past was unavailable, the database was constructed along the lines of areal interpolation, except that the population numbers were calculated using a spreadsheet instead of spatial attributes. In practice, the calculation for temporally comparable municipal populations can be expressed as follows.

$$P_{t-\Delta t} = p_{t-\Delta t}\left(1 + \left(\sum_{i=1}^{n}\frac{p_{part_{t_i}}}{p_{tot_t}}\right) - \left(\sum_{j=1}^{n}\frac{p_{div_{t_j}}}{p_{tot_t}}\right)\right) + \sum_{k=1}^{n} p_{merg_{t-\Delta t_k}} \tag{2}$$

where $P_{t-\Delta t}$ is the municipal population estimate of the point of time t-Δt fitted to the boundaries at the point of time *t*, $p_{t-\Delta t}$ is the original population amount of the municipality at point of time *t*-Δt, $ptot_t$ is the total population of the municipality at the point of time *t*, $ppart_{t_i}$ is the added population of partial municipal merge at the point of time *t* when *i* is the number referring to mergers, and $pdiv_{t_j}$ is the subtracted population of partial municipal dissolution at the point of time *t* when *j* is the number referring to dissolutions, and $pmerg_{t-\Delta t_k}$ is the added population of the merged municipality at the point of time t-Δt when k is the number referring to the merged municipality. To construct a temporally consistent database covering several decades, the calculation had to be done for all decades and all municipalities affected by the structural changes.

The number of municipalities in 2007 was 416 and the exact population numbers were available for 206 municipalities having a stable structure from 1880 to 2007. Changes in municipal boundaries without explicitly reported spatial arrangements had to be excluded from the calculations because of missing or inadequate data. The spatial reference for the municipalities in the accessibility calculations was the estimated location of the main centres.

Population grid databases from Statistics Finland (see Grid database 2010) were applied to calculate the dependent variable, population change at six grid cell resolutions (side lengths of 2, 4, 8, 12, 16 and 24 km) (papers III and IV) and also at a built-up area scale (paper II). For grid-based analysis, a 2×2 km resolution was applied as the basic unit of calculations and for lower resolutions, population weighted averages were computed. For the purposes of this thesis, the grid cell-based population data was obtained for the years 1970, 1980, 1990, 2000, 2003, 2007, 2008 and 2009, which was the most recent available data at the time of the last empirical analysis of the study. The original resolution of grid data was 1×1 km, except resolution of 2008 data was 250×250 m. In accessibility calculations, each grid cell was represented by its centroid location. The population grid data is maintained by Statistics Finland and it is based on population registers. This means that instead of simulated population probabilities, real-word residential information from the last day of the year is used, which remarkably improves the reliability of the data.

The population grid database (2004) was also used in constructing explanatory variables representing socio-economic differences in paper IV. The grid cell database consists of 108 records, in categories of age structure, level of education, consumer structure of the population, size and household life stage, household consumer structure, buildings and housing, workplaces and the main type of activity of the population from the years 2001–2003. These records were applied in constructing an extensive set of 58 socio-economic variables. After analysing the potential effect of these variables on population change, four of them were selected as final models. The variables for each scale of analysis were built by aggregating the data into the larger resolutions. Due to the protection of privacy, some database records were unavailable in grid cells containing fewer than 10 inhabitants, the cells of which had to be excluded from the analysis. Thus, the remaining data does not cover sparsely populated deep peripheries. However, 91.9 % of the Finnish population is included in the analysis. Population grid data was also applied in computing the grid cell population density.

Built-up areas are used as an alternative to administrative municipal units in paper II. The built-up area level population change was computed by utilising the 1×1 km population grid database. The boundaries are defined by Statistics Finland, on the basis of clusters of buildings including at least 200 inhabitants, where the distance between the buildings does not exceed 200 m. Although the areas were kept constant during each of the assessed decades, they differ between the decades as a result of the changes in the distribution of the population. Gravity centroids were applied as the spatial reference for built-up areas, except in the case of the capital Helsinki, for which the most populated grid cell was used.

5.2 Methods

In this study, the fastest route travel times between origins and destinations and the potential accessibility indicators based on them were applied to assess accessibility. Accessibility computations were carried out by using GIS. Due to explorative characteristics of this study, non-linear multiple regression generalized additive models (GAMs) were used to establish the relationships between dependent and explanatory variables.

5.2.1 GIS-based accessibility analyses

The accessibility calculations and related indicators applied in this study are based on the estimated fastest routes and travel times between origin and destination locations via a GIS-based model of the transport network. Estimating the fastest travel time in this thesis is based on a non-planar graph, in which vertices represent the members (*e.g.* roads) of the set and edges represent the connections among those members (*e.g.* crossroads) and the fastest route travel time computations are carried out with the classical Dijkstra algorithm (Dijkstra 1959). The GIS methods applied in the fastest route estimation are included in Esri ArcGIS 9.2, 9.3 and 10.0 Network Analyst (ArcGIS Network Analyst 2012). Python 3.0 and newer versions were applied in needed programming or scripting (Python Programming Language 2012).

Rail and air and transport accessibility variables were calculated as road network-based travel times (see Rietveld and Bruinsma, 1998, 34–35) from each municipal centre, built-up area gravity centroid or grid cell centroid to the nearest railway station (papers II, III and IV) and airport (papers III and IV), whereas in paper I, the distance to nearest railway station was used. This highly simplified approach enabled a statistical analysis of the effect of road-based accessibility potentials and other transport modes simultaneously in a multiple regression. Other accessibility indicators could be applied for this task as well. For example, multimodal accessibility potentials were considered for railway accessibility transports in papers I, II, and III, but due to problems related to multicollinearity (see Brauner and Shacham 1998; MacNally 2000), this approach had to be abandoned. In addition, in paper II, the accessibility to a railway network was also measured as a direct distance and as a binary variable indicating the presence of a railway network access. A wider selection of indicators would also increase the number of models remarkably. Hence, the clearest indicator, travel time, was selected. In papers III and IV, the population weighted averages of 2×2 km computations were used for larger scales.

To reasonably model road-based accessibility on several scales, a gravity-based potential accessibility analysis was applied (see Harris 1954; Bruinsma and Rietveld 1993; Geertman and Ritsema van Eck 1995; Youshida and Deichmann 2009; Gutiérrez *et al.* 2010). Potential accessibility is a measure relating the centrality and peripherality of locations and it has been widely applied in urban and geographical studies since the late 1940s (Geurs and van Wee 2004: 133–134). By potential accessibility, the clustered or isolated population concentrations and intermediate or deep peripheries can be differentiated numerically using an attracting attribute, like population, and connectivity by the transport network involving a friction measure, like time or cost. Potential accessibility has been successfully used in explaining the distribution patterns of populations (Song 1996). It lends itself well to place accessibility, the approach of which was analysed in this thesis, but has remarkable shortcomings if applied to individuals (Kwan 1998).

Potential accessibility $\mathbf{a} = (a_1, a_2, a_3 \ldots a_n)$ can be calculated for a location by dividing the population of all other locations by the travel (time-) distance separating the location and each of the other locations, and summarising these values. Depending on the characteristics of the transport, behaviour and analysed area, the computation is based on different types of distance decay functions. The linear and power functions are common types of distance decays (equation 3), but more often accessibility potentials $\mathbf{b} = (b_1, b_2, b_3 \ldots b_n)$ are based on negative exponential functions (equation 4):

$$\mathbf{a} = \sum_{j=1}^{n} \frac{P_j}{d_{ij}^{\alpha}} \quad \mathbf{a} = [\mathbf{a_i}]_{n\times 1} \tag{3}$$

$$\mathbf{b} = \sum_{j=1}^{n} \frac{P_j}{\mathrm{e}^{-\beta d_{ij}}} \quad \mathbf{b} = [\mathbf{b_i}]_{n\times 1} \tag{4}$$

where a and b are the potential accessibility vectors, d_{ij} is the distance between the location i and j, P_j is the population of the related destination location, n is the number of origins and destinations, and a and β are parameters for transport friction, indicating the efficiency of the transport system and the interest in moving and $e \approx 2{,}71828$.

Parameters a and β are highly dependent on the type of activity modelled. A greater distinction between nearby and distant destinations can be achieved by increasing a or decreasing β. At a local level, more extreme distance decay functions are used and on a large scale the distance decay is more gradual. In paper I, linear (a=1)distance decay function is used (see *e.g.* Gutiérrez 2001; Holl 2007). In paper II and III, quadratic (a=2) distance decay was also applied. In paper IV, four negative exponential functions were applied. The most gradual function (β=0.005), is associated with the European scale analysis by Spiekermann and Wegener (2007), while steeper functions (β=0.05) and (β=0.02) have been applied by Andersson and Karlsson (2007) for extra-regional accessibility and for intra-municipal accessibility. The steepest negative exponential function (β=0.0946) was estimated on the basis of trip survey data (Kalenoja and Kiiskilä 2010) in paper IV. The empirically estimated parameter corresponds well to function β=0.1, used by Andersson and Karlsson (2007) for intra-regional accessibility. Applied distance decay functions and associated travel frequencies are presented in Figure 3. As the regression model uses only relative differences between observations, not absolute, the numbers are scaled to one to facilitate comparison.

As its name implies, potential accessibility reflects the opportunity to access some particular objects. The problem of self-potential is often encountered in the calculation of the potential accessibility to large areas, because a remarkable quantum of objects may be located in an area for which the potential will be calculated (see Geertman and Ritsema van Eck 1995). In other words, are the objects technically considered as fully, partly or not accessed if they are located to the area already? Solutions to manage the problem are, *e.g.,* population-based self-potential estimates (Gutiérrez 2001) and Euclidian space-based approximations (Frost and Spence 1995). In order to include self-potential in the analysis, the potential accessibility $\mathbf{c} = (\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3 \ldots \mathbf{c}_n)$ may be written as:

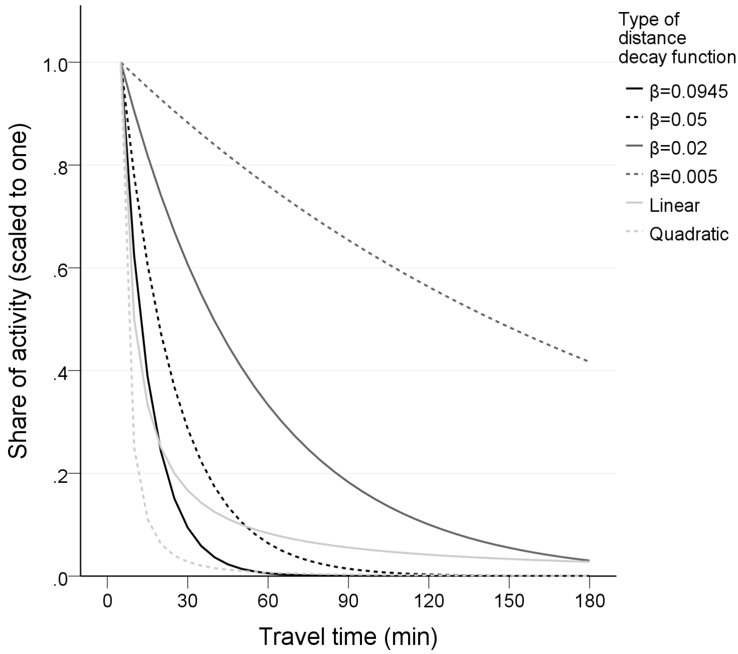

Figure 3. Applied distance decay functions in estimating the number of travels in relation to travel time. The maximum of each function is scaled to one.

$$\mathbf{c} = \frac{P_i}{d_{ii}^{\alpha}} + \sum_{j=1}^{n} \frac{P_j}{d_{ij}^{\alpha}} \quad \mathbf{c} = [\mathbf{c_i}]_{n\times 1} \tag{5}$$

where *Pi* is the population of area *i*, and d_{ii} is the mean internal distance of the area.

As there is no commonly accepted definition to describe the mean internal distance or travel time for the estimation of self-potential, a method for estimation was developed in paper II. The internal distance d_{ii} is intended to be defined more precisely by using a ratio approximating the difference between the shape of the unit and a circle (paper II), and the formulation can be expressed as:

$$\mathrm{d_{ii}} = \left(\left(1 + \ln\left(\frac{C/2\pi}{\sqrt{(A/\pi)}}\right)\right) * \frac{\sqrt{(A/\pi)}}{\sqrt{2}}\right) / \mathrm{v} \tag{6}$$

where C is the perimeter, A is the area as square kilometres, and v is the speed as kilometres per hour. This internal distance approach, of course, has some shortcomings, such as the ignoring of population distribution and congestion, but it enhances the estimate compared to purely population-based estimates (see *e.g.* Gutiérrez 2001: 236).

In accurate resolution grid cells, self-potential is commensurate mainly to population density. In addition, self-potential has little significance to the potential at accurate resolution computations. In papers III and IV, population density is used as an explanatory variable and, consequently, the self-potential was not included in the potentials. In these papers, potential accessibility to more coarse resolutions was calculated as the population-weighted averages of potentials at 2×2 km resolution, which provided a more realistic outcome than applying self-potential estimates. In paper I, no self-potential parameter is estimated, but the municipal population density is applied.

5.2.2 Statistical analysis, generalized additive models (GAM)

The relationships between population change and explanatory variables were scrutinized by non-linear and non-monotonic regression analysis, generalized additive models (GAMs) (Hastie and Tibshirani 1990). GAMs afford a great deal of freedom to specify the model for a different setting and a quantity of statistical mathematics is included in the structure of model. Even though it is not possible to thoroughly review GAMs in this thesis, a general overview is provided. When compared to other advanced modelling techniques GAMs have been proven to be effective *in both explaining and predicting spatial patterns* in the context of physical geography (*e.g.* Marmion *et al.* 2008). The main benefit of GAMs for this study, in comparison to conventional multiple linear regression, is that they rely more on the data-driven functional form than a model-driven form (Yee and Mitchell 1991). Classical regression approaches are based on a predetermined function type, such as linear, logarithmic or exponential, which is to be fitted in relation to the response variable with a set of explanatory variables. GAMs allow the capturing of the relationships with smooth functions which allow the identification of more complicated non-parametric relationships with a minimal loss of information and models. GAM may be expressed as:

$$g(\mu_i) = \mathbf{X}_i^* \boldsymbol{\theta} + f1(x_{1i}) + f2(x_{2i}) + f3(x_{3i}) + \ldots \quad (7)$$

where $\mu_i \equiv (Y_i)$ and Y_i is the response variable; is the link function; if strictly parametric model components exist (which are not applied in the models of this thesis) X^*_i is the row of the model matrix for model components and θ is the corresponding vector; x_k explanatory variable candidates and fj are smooth functions. In this study, the identity link function $f(x) = x$ is applied. Smooth functions are based on cubic splines, established with a conservative four degrees of freedom, in order to allow some complexity in the models, but to avoid over-fitting of the data. Cubic spline is a curve constructed from sections of cubic polynomial joined together so that the curve is continuous up to the second derivative. The effects of different smoothers were tested with intuitive samples. In addition to more accurate model fits and a more stable confidence interval within outliers, more flexible smoothers increased explanatory power the model (Wood 2006).

GAM models for all the papers were constructed with S-Plus 6.1 and a method for automated GAMs construction, Generalized Regression Analysis and Spatial Prediction (GRASP) (Lehmann *et al.* 2002) was applied for GAM building in paper IV. In addition to response curves, explained deviance, alone contribution and drop contribution values and confidence intervals of 95 percent and *p*-values were also used to support the analysis. The functionality of different types of model constellations in explaining population change and particular variables aimed to be included in models was considered by explained deviance. Alone contribution means an explained deviance of a univariate model and it refers to the potential importance of each variable. Drop contribution refers to the importance of each variable within the full model. Drop contributions are obtained by dropping each particular explanatory variable from the multivariate model and by calculating the associated change in deviance (Lehmann *et al.* 2005: 42). In other words, variables having a high drop contribution can explain a considerable amount of deviance that other variables are unable to explain. The predictive ability of GAMs was validated in papers III and IV by applying full datasets with all observations or four times cross-validation. Three stages can be defined in the validation process. First, establishing the relationship between dependent and explanatory variables to calibrate models. Second, the use of calibrated models to predict population change. Third, measuring the correlation between population change and predictions in each case. In papers I, II and III, Spearman Rho correlation factors were applied to avoid any bias caused by outliers, whereas Pearson correlation is built into GRASP, which was applied in paper IV.

6. Review of the results

The five research questions of this thesis are answered in this chapter. Whereas the original research papers contribute to answering many research questions, each sub-chapter will answer each of the particular questions as numbered in chapter 3. First, the relationship between road-based potential accessibility and population concentrations are considered. Second, the effect of the type and steepness of distance decay function (applied in computing potential accessibility) to this relationship is assessed. Third, the matter of scale in the consistency of the models is dissected. Fourth, the effects of rail and air transport facilities, key socio-economic variables and population density to population change are expounded. Finally, the predictive ability of established models is considered. The variable types consisting of the final models explaining population change in each paper are compiled in relation to scale and time and presented in Figure 4.

6.1 Accessibility and urbanisation

In general, road network-based potential accessibility captured the attractiveness of cities and population concentrations well in the statistical models explaining population change. By this approach, the concentration of the Finnish population has been evident at a municipal level since the end of the 19th century. The intensity of concentration has increased in conjunction with structural changes and related growth of the economy and the consequent relocation of the population. The growth of modernising cities intensified during the 1920s, characterized by strong economic growth and industrialisation, and diminished by the recession of the 1930s. After the post-war relocation of the population,

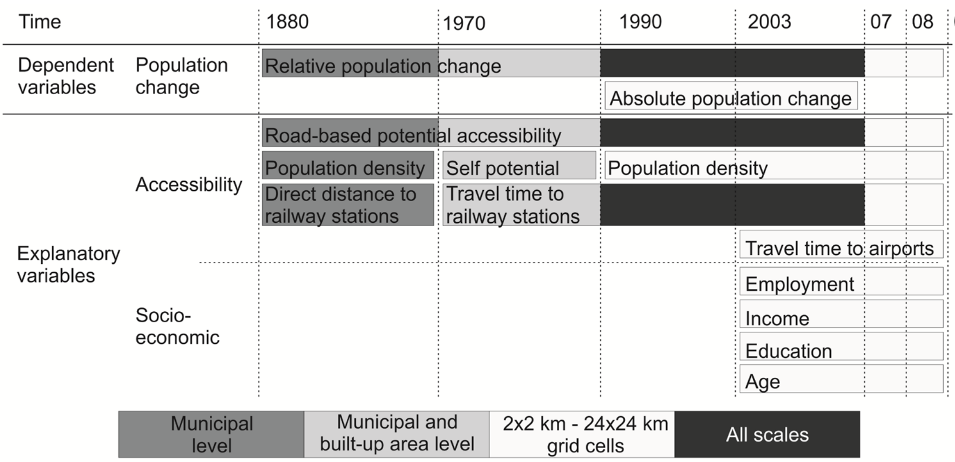

Figure 4. Dependent and explanatory variables selected to final models in relation to scale and time.

the concentration process intensified again and culminated during the rapid urbanisation of the 1960s (paper I). During the 1970s and 1980s, the concentration diminished, but during the 1990s, urban centres and clusters intensively attracted a population again, the trend which continues to the present date (papers II and III). This return might be attributable to the major changes in Finnish society, such as the opening of the economy and the shift towards developing centres instead of subsidising peripheries.

In addition to regional concentration, urban sprawl, a dispersive trend of a growing urban edge was prevailing in models considering population change after 1970. During 1970–1990, population change in built-up areas could not be explained like the change at a municipal level. The finding can be traced to be mainly scale and data related, since the population growth in municipalities has occurred at the edge of the defined built-up areas, in addition to centres, and simultaneously population in the peripheries has concentrated at a local level towards peripheral town centres, whilst the surrounding areas have degenerated rapidly (paper II). The trend is also visible in grid-based models. With fine resolution models, it could be noticed that the population concentration in a centre-periphery axis stagnates in the suburbs. Models also showed a lesser population growth at easy accessible locations with high population densities (papers II and III). This is partly related to the mathematical relation of population change and population density as a variable. However, the mechanics of defining multiple regression by GAM, with explanatory variables having some mutual correlation, may also affect inconstancies. However, the trend is visible in models carried out with relative and absolute population change.

From an analytical point of view, accessibility potentials consist of population distribution and the capacity of the road network. Thus, it is not possible to exactly define the importance of these components to accessibility potentials with this study setting. Population is, however, a much more varying value in comparison to travel speeds and network geometry, which refers to the concept that the population component has a more essential role.

6.2 Friction of distance in reflecting urban structure

A highly detailed view of the urban structure is achieved, when the steep distance decay functions were used in computing accessibility potentials. Two different types of distance decay functions, linear and quadratic, were tested in papers I, II and III and in paper IV, four different negative exponential distance decay functions were added to the analysis. In the time before motorisation, relatively gradual, linear distance decay-based accessibility potentials succeeded in being connected to population change with corresponding explanatory performance compared to steeper, quadratic, distance decay function-based potentials. The linear distance decay function-based accessibility potentials were selected for the final models, due to their simplicity and comparability. As these potentials rather reflect a regional accessibility, intra-municipal accessibility is considered through a population density variable.

After motorisation, accessibility potentials with steep distance decay functions succeeded in predicting population change better than potentials with gradual functions. A quadratic function was more suitable than one that was linear (papers II and III). Nevertheless, an empirically estimated negative exponential distance decay function (β=0.0945) proved to form potentials that predict population change most efficiently (paper IV). The form of this distance decay functions is, however, very close to the quadratic function. Thus the predictive ability of the potentials based on quadratic function succeeded in explaining population change nearly as efficiently. A possible explanation for the success of the steep functions is the accurate scales applied in the study. In contrast to the use of administrative regions in the calculation of potential accessibility, requiring the use of internal distance estimates and self-potential, high resolution grids make it possible to detect short-distance patterns. Since a greatly detailed view of an urban structure is achieved with the relatively steep distance decay functions with a high predictive ability towards population change, the applying of steep functions with accurate resolution grid cell data is supported. However, the potentials of larger spatial units relies more on self-potentials. In accurate resolutions, the extent of this component is dominated by the short distance components of the potential. Hence, it may be advisable that applying the steep functions to the potentials of large areas are done with caution.

6.3 Matter of scale in the consistency of models

The major trend prevailing in models explaining population change at different scales is that the explanatory power and predictive performance decreases with the scale becoming more accurate. Furthermore, with a careful selection of accessibility and socio-economic variables, population change can be modelled tolerably at accurate resolutions and well at coarse resolutions. Models at a municipal level showed remarkably better explanatory performance than built-up area level models (paper II). The trend was unambiguous with models that were carried out with six different resolution grid cell datasets, from 2×2 km to 24×24 km (papers III and IV). Validations considering the predictive ability of models showed corresponding results. With full datasets, predictions were slightly better than when applying the average results of four times cross-validation and larger resolution predictions were more accurate than smaller resolution ones (paper IV). In paper III, it was tested that this disparity cannot be explained by the huge variation in the number of observations, as the trend corresponds in models carried out with a constant size, random selection of observations. This suggests that the consistency of models is adequate for spatial extrapolations.

The theoretical and statistical relevance of explanatory variables at different scales were considered before establishing final models for papers II, III and IV. Road-based potential accessibility, airport accessibility and also key socio-economic variables maintained the characteristics of their relationship between population change, regardless of the varying size of spatial units. On this basis, it can be concluded that with larger spatial units, stochasticity related to local variation and an amount of individual decisions gets averaged. This averaged view forms an areal-level pattern, which is possible to be captured by a model with explanatory or predictive purposes.

6.4 Effects of other variables

In this chapter, the effect of other variables included in multivariate models is considered. First, the effects of rail and air accessibility are assessed. Second, the relationship between key socio-economic variables, *i.e.* average incomes, the share of academic education, employment rate and average mean age, and population density is contemplated. Third, the different roles of a population density variable in explaining population change with other variables are considered.

6.4.1 Effect of rail and air transport facilities

A dominating population change process, concentration towards urban areas, needed to be modelled accurately enough with the road-based potential accessibility indicator, before the disentangling of the effect of other transport modes was possible. In multiple regression analysis, the effect of rail and air transport facility accessibility was analysed in conjunction with other explanatory variables. In addition to the travel time to the nearest station, other tested approaches were multimodal potential accessibility (papers I and II), direct distance and a binary variable related to the presence of a railway station in the area (paper II). Multimodal potential accessibility was calculated by using the fastest travel times on the road and railway networks connected by railway stations. However, it did not improve the explanatory power of the models and problems related to multicollinearity occurred. The travel time to the nearest station was selected as an accessibility variable, because direct distance and a binary variable were completely insignificant in statistical terms in joint GAMs, which underlines the importance of a network-based travel time measurement, instead of direct distance or buffering, when analysing societal phenomena (paper II).

The importance of different transport modes in explaining population change was compared and road network-based potential accessibility had the most essential role in the great majority of cases throughout time (papers I and II) and scales (papers III and IV). The expansion of the railway transport system was nearly completed at the turn of the 20th century. After the turbulent beginning of the century, diminishing out-migration and economic growth, railway accessibility was noted to have a clear effect on population change at a municipal level during the 1920s. After the recession of the 1930s, the effect of railway accessibility was clear during the period 1940–1970 (paper I). Again, during the 1970s, the effect of railway accessibility diminished. The decade was characterized by increasing motorisation, but railways, however, operated notably at a local level. After remarkable investments in long-haul transport, the slight effect of railway accessibility could be captured in 2000–2007 (paper II).

Even though municipal-level models indicated that railway accessibility has an effect on population change, the accessibility to railway stations proved to be statistically insignificant in grid-based models and in a few models where railway accessibility had a minor statistical significance, theoretically inconsistent fits were captured. Consequently, the railway accessibility variable had to be excluded from the final models in papers III and IV. A reason for this may be the small (approximately 5–7 %) share of railway travel after motorisation.

The effect of airport accessibility was tested with grid cell-based models in papers III and IV. Airport accessibility was remarkably related to population change. Accessibility to airports was noted to have a clear positive effect on the population, starting in the immediate proximity of airports. This finding clearly underpins the need for active airports in the regions striving to develop. Nonetheless, it is important to scrutinize the actual characteristics of causality behind this relationship before making further conclusions. The relationship was highly consistent in all the resolutions, but the explanatory power of airport accessibility increased with scale (papers III and IV).

6.4.2 Effect of key socio-economic variables

With the aim of testing the effect of other relevant variables in explaining population change, and to also reduce the risk of the omitted variables problem, a carefully selected set of key socio-economic variables were included in the models in paper IV. The analysis resulted in three important findings: First, based on the drop contributions, urbanisation was a process that was captured very efficiently by a potential accessibility variable and it was clearly the most essential trend explaining population change during the intensifying period of motorisation. Second, socio-economic variables produced significantly stable relationships between population change in relation to resolution and the explanatory performance power of each particular variable also increases along with the resolution. Average incomes and the higher share of academic education are mainly related to population growth. Surprisingly, employment has only a slight effect on population growth. The age variable clearly reflects the natural population increase and agglomeration of young people in urban areas. However, an exceptional growth trend can be found in peripheral grid cells characterized by the mean age of over 50. However, any possible association of this trend towards remigration or second housing is an open question. Third, socio-economic variables together largely explain much of the deviance that cannot be explained with accessibility variables. Hence, with accessibility and population density variables, it is possible to make predictions of population change at a grid cell basis, but the performance of the model may be enhanced, to some extent, with socio-economic variables.

6.4.3 Different roles of population density variable

Population density is used widely in defining different areas on the centre-periphery axis. As models in all the papers show, it was possible to model population change to some extent with a population density variable. Indeed, a population concentrated in already urbanized regions is characterized by high population densities. The population density variable was important to be included in the models for three reasons. First, for the credibility of the setting of this study, it was important to compare the performance of the population density and potential accessibility variables. On the basis of the drop contribution, before motorisation, the population density explains population change better or equally as good as potential accessibility (paper I). For municipal and built-up area level models in paper II, internal distance and self-potential estimates were used and the population density was excluded. Second, due to a variation in the population density in grid cells, the same absolute population increase in adjacent grid cells may vary greatly when presented as a relative number. Population density explains this disparity when the population density variable is used in models that explain a relative population change. However, the population density does not capture the actual population change to the same extent in GAMs than the potential accessibility does (papers III and IV). Third, population density also reflects the high residential costs and land-use intensity in areas with high accessibility. In modelling an absolute population change, this role is clearly evident (paper III).

6.5 Extrapolation of models

In addition to considering the predictive performance at space, temporal predictions were also evaluated (paper III). Population changes at different grid cell resolutions during the period 2000–2008 were predicted with models calibrated with datasets considering the datasets of the periods 2000–2008 and 1990–2000 (paper III). The predictive performance of models noted to be consistent also over time. An essential finding was that models with all observations and with small size random samples had a very similar predictive ability (paper III). This shows that the modelled patterns do not fragment much when the observation amount is decreased. However, for extrapolation, the predictive performance of the models is dependent on the representativeness of conditions in calibration datasets in relation to the prediction dataset and the characteristics of population dynamics have to remain the same in space or time.

7. Concluding discussion

The urbanisation process and sprawl, or more specifically expressed, population change at a centre-edge-periphery axis, can be perceived with a high specificity by applying a potential accessibility indicator in characterising the centrality and peripherality of locations in statistical models explaining population change. Potential accessibility, a ratio expressing the geographical distribution of transport infrastructure capacity and population, reflects the quantum of the activity and gravitation related to the interaction of the population. This relationship indicates the strong influence of scale economies on Finnish population change. Before the arrival of industry and the railroad, the population of Finland was almost evenly distributed. When the industrialising cities began to grow during the late nineteenth century, the potential accessibility of other municipalities began to matter. The heavy concentration of people in regional centres coincided with the major wave of urbanisation that took place during the 1960s, whereupon Finland rapidly motorized. These trends are associated with the widening scale of the Finnish economy, from the local to the regional level and then from the regional to the interregional level, covering the state.

When the major trend population, concentration, were explained, the lesser effects of other transport modes could be revealed more reliably. The accessibility of airports was included in the models since 1990 and had a clear growth effect on the surrounding population, whereas railway accessibility could be mainly related to urban growth between 1940–1970 and during the 1920s. It is interesting that the population density is much weaker in capturing population change in comparison to potential accessibility since the 1960s. Combining the effect of transport accessibility and socio-economic variables enabled the gain of new, or more accurate, insights into population change since 2003. It seems that population growth occurs in areas with high overall education levels, incomes and employment, whereas a high average age is usually associated with a decreasing population.

In the geographical analysis, the correctness of the presumptions of scale are important, because when they are incorrect, analyses may produce quasi images of the actual patterns that are to be scrutinized. The relationship between transport accessibility and population change was relevant below the most accurate administrative division, the municipal level. The finding underpins the relevance of grid cell-based models in analysing accessibility. However, despite the effect of accessibility being visible with observations on an accurate scale, the actual pattern exists on a larger scale, for which the accessibility calculations should be generalized and development predictions should be made. Hence, also in ad hoc studies and related results, used in decision making and policy purposes, the scales of analyses should be relevant for phenomena, not data derivatives.

With accessibility, socio-economic and population density variables, it is possible to make predictions for a population change on a grid cell basis. Population change relates strongly to explanatory variables on a large scale because of a pattern created by the large number of individual choices. Nonetheless, the model involving these same choices on an excessively accurate scale maintains its characteristics, but loses much predictive ability. At an accurate scale, there is no actual pattern, due to increasing stochasticity related to micro-level properties, the outcomes of individual choices, which are not able to be significantly explained even if there are accurate socio-economic variables included. By improving data and accessibility indicators, population change on a very accurate scale could possibly be explained better. However, in modelling population change, there may be a certain scale threshold, which cannot be exceeded with quantitative analysis, even if all imaginable data were used. This study setting provides only a preliminary support to the applicability of the model for long-term predictions, an issue which clearly needs to be scrutinized within a larger study setting. As population change in Finland was previously modelled mainly through municipal division, grid cell-based models would improve the spatial accuracy remarkably. This can be considered to be particularly important for spatial planning, locating services and business activities and, more generally, for policy making, because a majority of investments are made for the long term, rather than for the present situation.

The results of this thesis have to be discussed in relation to the classical analytical problems of geography: the modifiable areal unit problem, ecological fallacy and spatial autocorrelation. It is definitely an important finding that models with selected variables produced consistent fits for different scales.Hence, a classical problem related to modifiable areal units does not seem to pertain to the relationships established for the resolutions of this study. However, an ecological fallacy needs to be seriously taken into account. The explanatory power and predictive ability of models decreases strongly when the scale of the analysis becomes more accurate. Correspondingly, the results of the study cannot be extrapolated to resolutions that are more accurate than those analysed, and particularly not to the individual level. Some spatial autocorrelation exists inevitably in any population data that consists of small areal units. In a society that is connected by transport networks, potential accessibility may also be operationalized to models with a motivation to reduce this bias. According to Andersson and Gråsjö (2005), the significance of the spatially discounted variables can be interpreted as spatial dependence, the presence of any kind of spatial dependence can invalidate regression results, and consequently, the autocorrelation cannot be ignored. However, their model shows that spatial dependence in the error terms vanishes when the model includes accessibility variables and, if accessibility variables are statistically significant, it suggests that spatial autocorrelation problems are significantly reduced.

There are many opportunities for the further development of potential accessibility methodology. The models had a high explanatory power in terms of population change, especially when the quadratic or steep negative exponential (β=0.0945) distance friction parameters were used in paper IV. Often, the distance decay parameter is selected arbitrarily or on the basis of surveys. Potential accessibility is based on the idea of gravity and travel densities, but the sphere of living and the matter of distance are, in the end, very specific to the individual. In an ideal case, the surveys could be developed to cover the particular activities of individuals with extensive socio-economic portraits: a good example of such a survey is the GPS-based TeleFOT (2012). Regional-, areal-, local- or even individual-level functions would possibly be defined and applied for more narrow studies or distance decay could be replaced with other types of functions (see *e.g.* Kwan 1998). The empirical function may be improved for further studies, if several regional traffic surveys can be combined or an adequately extensive national traffic survey is available. However, the empirical distance decay function improved the performance of models explaining and predicting population change, in comparison to literature-based functions and the quadratic distance decay function. A very simple and clear linear distance decay was selected for the pre-motorized era. The function allows for some error in the travel time estimation, because, instead of the absolute, only relative travel times affect the proportionality of the potentials.

In Finland, traveling has been clearly based on private cars with some exceptions since the 1970s. The supply and demand of transport are not homogenous and multi-modal travel chains are common in the largest Finnish cities. The data that is used in this study is of high quality in terms of geometric accuracy, spatial coverage and information quality. If there is a desire to improve accessibility models for future studies, the challenge is to take the required behavioural factors about common routines, such as choosing between transport modes in different situations, into account and route estimates could be enhanced to follow actual travelling patterns instead of the fastest routes. Accessibility could be considered in a multimodal context by including public transports in the analysis, a task that would be realistic for many study settings. However, the inclusion of such variables in this study would cause problems with multicollinearity or only a single, summarising, accessibility variable could be included. In addition, it is possible to estimate congestion factors and include them accurately in the analysis carried out at a local or regional scale. While improving the models with more precise data is a manageable task, constructing this type of datasets would require a substantial effort.

Studies often rely on regressions based on parametric response shapes and non-linear relationships are analysed through transformations. In addition to model descriptives, the information about the nature of the relation between accessibility and population change is highly important in explorative research. Although the established models produced response curves that were mostly increasing or decreasing, mainly linear forms covering response shapes were obtained infrequently. GAMs proved to be an effective methodology for revealing the characteristics of this relation quite explicitly, due to the data-driven form of relation. If traditional regression methods, such as linear regression, were used in this study, a great deal of information would be lost, since the actual relation is other than linear. Also, in the linear framework, either a large amount of observations should have been excluded or the characteristics of the relationship and explanatory power would have been harmed. Therefore, there were clear benefits in establishing the relationships in the explorative analysis in a non-linear form. However, in measuring the accessibility of rail and air transport, a highly simplified approach was used due to problems related to multicollinearity. If statistical tools handling multicollinearity will be developed and the high correlation between the different types of accessibility can be tolerated in statistical models, this will enable the more free use of accessibility indicators in similar study settings.

If there were accurate socio-economic variables available within the required levels of accuracy for the whole research period, the actual role of accessibility in explaining population change could be considered on a more certain basis. Currently, there is bigger risk of the omitted variables problem in papers I, II and III than in paper IV. In addition to reducing this problem, including more relevant variables in models improves also the predictive ability in general and, particularly, at the most accurate resolution, where the conditions in adjacent grid cells may vary substantially (paper IV). With respect to the availability of better data sources, some aspects could be developed. Due to the protection of the privacy of individuals, socio economic data at 1×1 km grid cells with small populations was not available for research use (paper IV). Thus, a relatively large proportion of deep peripheries was inevitably excluded from the models. The different level planning systems can be considered to have effects as well, but they are quite difficult to quantify. The locations of the actual sites of construction are decided in municipal-level planning and zoning, so also population flows on a local scale are dependent to planning. Thus, including planning related variables to models would improve their explanatory and predictive performance on built-up-area- and grid cell-level. At contrast, large-scale planning has only indirect effects on migration and population change.

In general, the results pertaining to the long-term concentrating effect of accessibility noticed in Finland are interesting in four ways. First, modelling population change affected by a strong concentration tendency on a large scale, especially related to the road transport system and population distribution, enables a more accurate spatial estimation of population change. This averaged view forms an areal pattern, which is able to be captured by a model with explanatory or predictive purposes. However, it is necessary to note that this areal pattern is not able to be related *per se* to individuals, as the phenomenon occurs in groups. Second, this prevailing process in a regional structure seems to improve the competitiveness of the centres and the state due to scale economics and also structural energy efficiency by shortening the distance related to compacting regional structure and decreasing transport needs. Third, taking this concentration trend into account in urban and regional planning and transport policies, the future transport lengths and emissions may be reduced due to condensing the spatial structure. Fourth, as only rare factors are solely local or aspatial, it can be stated that accessibility should be aimed towards bringing more into human geographical GIS studies and particularly to the models of population change. Still, regardless of how well the statistical models explain population change or other explained phenomenon, they rarely entail direct causality. Rather, the question is not so much about the accessibility itself, but the attracting phenomena (*e.g.* scale economies) prevailing within the accessibility.

The models and results of this thesis outline the status quo of Finnish population trends, and some policy-orientated questions can be discussed. Spatially and temporally consistent models show that over the last two decades, the population has concentrated towards areas at a regional scale, but population pressure exists at urban margins at a local scale. Nothing indicates that the trend will not continue. As it seems that, increasing accessibility mainly benefits the centres, it would be fair to ask if the policies supporting cohesion between centres and peripheries should be directed rather elsewhere than toward transport issues, as long as or until a decent level of transport networks are available in inhabited areas? Again, the proximity of airports seems to attract a population, which results probably from indirect location choices of companies and related employment and migration. Regardless of a remarkable rationalisation of railway transports and a huge decrease in the number peripheral stations, railway accessibility, unfortunately, does not have an observable effect on population change in general. When this trend is combined with a trend towards urban sprawl, it can be concluded that the potentials to enhance public transports at a local level by railways does not seem to be realistic, and the sprawl trend in general will impair opportunities to develop more efficient public transport system. Hence, the spread of urban structure should be limited. Another trend is a decrease in the population at deep peripheries. Of course, many exceptions to this can be found, but general trend is clear on the scale of the state. On the base of the results of this study, the growth of the peripheries in general does not seem to be realistic in the future, which finding should be applied in planning when there is a desire to make future scenarios on a realistic basis.

Finally, the significance of the analytical framework of this thesis is in exploring and revealing spatial trends and patterns that may be hidden to intuitive thinking and cartographical examination, rather than in defining direct causalities. Even though the results of the study are very country specific, the study setting and the modelling procedure that is used in this thesis can also be applied to other areas, if the GIS data for transport infrastructure, population and additionally socio-economic data are available at a required accuracy and sufficiently fine spatial resolution.

References

Ahvenainen, J. & A. Kuusenterä (1982). Teollisuus ja rakennustoiminta. *In* Ahvenainen, J., Pihkala, E. & V. Rasila (eds.): *Suomen taloushistoria,* 222–261. Kustannusosakeyhtiö Tammi, Helsinki.

AIP Suomi – Finland (2011) Finavia, Vantaa. Available at <https://ais.fi/ais/eaip/en/> Accessed August 2012

Alestalo, J. (1983). The concentration of population in Finland between 1880 and 1980. *Fennia* 161: 2, 263–288.

Alho, J. M. & B. D. Spencer (2005). *Statistical Demography and Forecasting. 410 p.* Springer, New York.

Alonso, W. (1964). *Location and land use.* Harvard university press, Cambridge.

Andersson, C., K. Frenken & A. Hellervik (2006). A complex network approach to urban growth. *Environment and Planning A* 38: 10, 1941–1964.

Andersson, M. & U. Gråsjö (2005). On the specification of regression models with spatial dependence – an application of the accessibility concept. *CESIS Electronic working paper series* 51. 27 p.

Andersson, M. & C. Karlsson (2007). Knowledge in regional economic growth – the role of knowledge accessibility. *Industry and innovation* 14: 2, 129–149.

ArcGIS Network Analyst (2012). Esri. Available at <http://www.esri.com/software/arcgis/extensions/networkanalyst> Accessed August 2012

Aro, T. (2007). Julkinen valta ja maassamuuttoa edistävät ja rajoittavat tekijät Suomessa 1880-luvulta 2000-luvulle. *Koulutussosiologian tutkimuskeskuksen raportti* 69. 324 p.

Banister, D. & Y. Berechman (2001). Transport investment and the promotion of economic growth. *Journal of Transport Geography* 9: 3, 209–218.

Biehl, D. (1991). The role of infrastructure in regional development. *In* Vickerman, R. W. *Infrastructure and Regional Development,* 9–35. Pion, London.

Black, W. R. (2003). *Transportation.* 375 p. The Guildford press, New York.

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* 22: 3, 547–581.

Brauner, N. & M. Shacham (1998). Role of range and precision of the independent variable in regression of data. *AIChE Journal* 44: 3, 603–611.

Brown, D., G. Elmes, K. Kemp, S. Macey & D. Mark (2004). Geographic Information Systems. *In* G. Gaile & C. Willmott (eds.): *Geography in America at the Dawn of the 21st Century*, 353–375. Oxford University Press, New York.

Brueckner, J. (2000). Urban Sprawl: Diagnosis and Remedies. *International Regional Science Review* 23: 2, 160–171.

Bruinsma, F. & P. Rietveld (1993). Urban agglomerations in European infrastructure networks. *Urban Studies* 30: 6, 919–934.

CAA's Air traffic statistics (2004). CAA Finland. Vantaa.

Chen, A., C. Yang, S. Kongsomsaksakul & M. Lee (2007). Network-based Accessibility Measures for Vulnerability Analysis of Degradable Transportation Networks. *Networks and Spatial Economics* 7: 3, 241–256.

Chi, G. (2010a). The impacts of highway expansion on population change: an integrated spatial approach. *Rural Sociology* 75: 1, 58–89.

Chi, G. (2010b). The Impacts of Transport Accessibility on Population Change across Rural, Suburban and Urban Areas: A Case Study of Wisconsin at Sub-county Levels. *Urban studies.* 49: 12, 2711–2731.

de Bok M. & F. van Oort. (2011). Agglomeration economies, accessibility, and the spatial choice behavior of relocating firms. *Journal of transport and Land use* 4: 1, 5–24.

de Smith M. J., M. F. Goodchild, & P. A. Longley (2007). *Geospatial Analysis, A comprehensive Guide to Principles, Techniques and Software Tools.* 2nd ed. The Winchelsea Press/Matador, Leicester.

Digiroad (2012). Finnish transport agency, Helsinki. Available at <http://www.digiroad.fi/dokumentit/en_GB/documents/_files/84853526732145072/default/Description_of_Digiroad_Data_Objects_31.pdf> Accessed August 2012

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

European Commission (2011). *Roadmap to a single European transport area: towards a competitive and resource efficient transport system.* White Paper.
Finavia's Air traffic statistics (2009). Finavia, Vantaa.
Finavia Oyj Vuosikertomus (2011). Finavia, Vantaa.
Finnish railway statistics (2010). Finnish transport agency 06/2010. Helsinki.
Fotheringham, A. S., C., Brunsdon & M. Charlton (2000). *Quantitative Geography.* 270 p. SAGE publications, London.
Frost, M. & N. Spence (1995). The rediscovery of accessibility and economic potential: the critical issue of self-potential. *Environment and Planning A* 27: 11, 1833–1848.
Fujita, M. & J.-F. Thisse (1996). Economics of Agglomeration. *Journal of the Japanese and international economies* 10: 4, 339–378.
Fujita, M., P. Krugman & A. J. Venables (1999). *The Spatial Economy*. The MIT Press, London.
Geertman S & J. Ritsema van Eck (1995). GIS and models of accessibility potential: an application in planning. *International Journal of Geographical Information Systems* 9: 1, 67–80.
Geurs, K. & J. Ritsema van Eck (2001). Accessibility measures: review and applications. Evaluation of accessibility impacts of land-use transport scenarios, and related social and economic impacts. *RIVM report* 408505 006. 265 p. Available at <http://www.rivm.nl/bibliotheek/rapporten/408505006.pdf> Accessed August 2012
Geurs, K. T. & B. and van Wee (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography* 12: 2, 127–140.
Gibson, C., E. Ostrom, & T. Ahn (2000). The concept of scale and the human dimensions of global change: a survey. *Ecological Economics* 32: 2, 217–239.
Glaeser, E. (1994). Cities information and economic growth. *Cityscape* 1: 1, 9–48.
Glaeser, E. (1998). Are cities dying? *The journal of economic perspectives*. 12: 2, 139–160.
Glaeser, E. L. & M. E. Kahn (2004). Sprawl and urban growth. *In* Henderson J. V. & J.-F. Thisse (eds.): *Handbook of regional and urban economics, cities and geography*, 2481–2527. Elsevier, Amsterdam.
Gløersen, E. (2009). Strong, specific and promising. *Nordregio Report* 2009: 2. 78 p.
Goodchild, M. (2004). GIScience, Geography, Form, and Process. *Annals of the Association of American Geographers* 94: 4, 709–714.
Graham, D. (2007). Agglomeration, productivity and transport investment. *Journal of Transport Economics and Policy* 41: 3, 317–343.
Grid database (1970). Statistics Finland, Helsinki.
Grid database (2004). Statistics Finland, Helsinki.
Grid database (2010). Statistic Finland, Helsinki. Available at <http://www.stat.fi/tup/ruututietokanta/index_en.html> Accessed August 2012
Guisan, A. & N. E. Zimmermann (2000). Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 2–3, 147–186.
Gutiérrez, J. (2001). Location, economic potential and daily accessibility: an analysis of the accessibility impact of the high-speed line Madrid–Barcelona–French border. *Journal of Transport Geography* 9: 4, 229–242.
Gutiérrez, J., A. Condeço-Melhorado & J. Martín (2010). Using accessibility indicators and GIS to assess spatial spillovers of transport infrastructure investment. *Journal of Transport Geography* 18: 1, 141–152.
Hakala, J. (1973). The availability of central-place services and the network of central places in Finland. *Nordia* 2. 28p.
Halme, T (1999). Muuttuva alue- ja yhdyskuntarakenne. *Nordia geographical publications* 28: 1. 150 p.
Harris, C. (1954). The market as a factor in the localization of industry in the United States. *Annals of the Association of American Geographers* 44: 4, 315–348.
Hastie, T. & R. Tibshirani (1990). *Generalized Additive Models.* 335 p. Chapman and Hall, London.
Haugen, K. (2011). The Advantage of 'Near': Which Accessibilities Matter to Whom? *European Journal of Transport and Infrastructure Research.* 11: 4, 368–388.
Heikkilä, E. (2003). Differential urbanisation in Finland. *Tijdschrift voor Economische en Sociale Geografie* 94: 1, 49–63.
Heikkilä, E. & M. Pikkarainen (2010). Differential population development in the regions of Finland. *Population, Space Place* 16: 4, 323–334.

Helminen, V., H. Rita, M. Ristimäki & P. Kontio (2012). Commuting to the centre in different urban structures. *Environment and Planning B: Planning and Design* 39: 2, 247–261.
Herzog, H. & D. Bjornstad (1982). Urbanization, interregional accessibility, and the decision to migrate. *Growth & Change* 13: 3, 21–25.
Holl, A. (2007). Twenty years of accessibility improvements. The case of the Spanish motorway building programme. *Journal of Transport Geography* 15: 4, 286–297.
Johansson, B., J. Klaesson & M. Olsson (2002). Time distances and labor market integration. *Papers in Regional Science* 81: 3, 305–327.
Kalenoja, H. & K. Kiiskilä (2010). *Oulu region travel survey 2009.* Oulun seutu, North Ostrobotnia Centre for Economic Development, Transport and the Environment, Ministry of Transport and Communications. 31 p. Available at <http://oulu.ouka.fi/seutu/pdf/OuluLiikennetutkimus2009.pdf> Accessed August 2012
Kalliokoski, A. (2003). Liikenteen hallinnan keinot ja vaikutukset ruuhka-aikoina. *Tiehallinnon selvityksiä* 16/2003. 101 p. Available at <http://virtual.vtt.fi/virtual/proj6/fits/julkaisut/VIKING_Y2002_D3_5S_1_Liha_keinot_vaikutukset_ruuhka.pdf> Accessed August 2012
Kauppinen, J. (2000). Muuttoliike Suomessa vuosina 1989−1994 koordinaattipohjaisten paikkatietojen perusteella. *Siirtolaisinstituutin tutkimuksia* A 22. 155 p.
Kilpailukykyä ja hyvinvointia vastuullisella liikenteellä. Valtioneuvoston liikennepoliittinen selonteko eduskunnalle 2012 (2012). Liikenne- ja viestintäministeriö. Liikennepolitiikan osasto. Liikennealan strategia. Available at <http://www.hare.vn.fi/upload/Julkaisut/17748/670671812912207.PDF> Accessed August 2012
Klaesson, J. & B. Johansson (2008). Agglomeration dynamics of business services. *CESIS Electronic Working Paper Series* 153, 20 p.
Koskinen, S., T. Martelin, I.-L. Notkola, V. Notkola, K. Pitkänen, M. Jalovaara, E. Mäenpää, A. Ruokolainen, M. Ryynänen & I. Söderling (2007). *Suomen väestö.* 351 p. Gaudeamus, Helsinki University Press, Helsinki.
Krugman, P. (1991). Increasing returns and economic geography. *The Journal of Political Economy* 99: 3, 483–499.
Kwan, M.-P (1998). Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework. *Geographical Analysis* 30: 3, 191–216.
Lehmann, A., J. Overton & J. Leathwick (2002). GRASP: Generalized regression analysis and spatial predictions. *Ecological Modelling* 157: 1–2, 189–207.
Lehmann, A., J. Overton & J. Leathwick (2005). *GRASP v.3.0 User's Manual.* Swiss Centre for Faunal Cartography, Switzerland. Available at <http://www.unige.ch/ia/climate/grasp/downloads/GRASP_ME_v3.2.pdf> Accessed August 2012
Lehtonen, O. & M. Tykkyläinen (2012). Työpaikkakehityksen alueelliset kehitysprosessit Itä-Suomessa 1994–2003.*Terra* 124: 2, 85–103.
Liu, W. & K. Seto (2008). Using the ART-MMAP neural network to model and predict urban growth: a spatiotemporal data mining approach. *Environment and Planning B: Planning and Design* 35: 2, 296–317.
Longley, P., M. Goodchild, D. Maguire & D. Rhind (2005). *Geographical information systems and science.* 2nd ed. John Wiley & Sons, Chichester.
MacKinnon, D., G. Pirie & M. Gather (2008). Transport and economic development *In* Knowles, R., J. Shaw & I. Docherty (eds.): *Transport Geographies: Mobilities, Flows and Spaces*, 10–28. Blackwell Publishing, Oxford.
MacNally, R. (2000). Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of – 'predictive' and 'explanatory' models. *Biodiversity and Conservation* 9: 5, 655–671.
Malpezzi, S. (2011). Urban Growth and development at Six Scales *In* Birch E. & S. Wachter (eds.): *Global Urbanization*, 48–66. University of Pennsylvania Press, Philadelphia.
Marceau, D. (1999). The scale issue in social and natural sciences. *Canadian Journal of Remote Sensing* 25: 4, 347–356.
Marmion, M., J. Hjort, W. Thuiller & M. Luoto (2008). A comparison of predictive methods in modeling the distribution of periglacial landforms in Finnish Lapland. *Earth Surface Processes and Landforms* 33: 14, 2241–2254.
Meentemeyer, V. (1989). Geographical perspectives of space, time and scale. *Landscape Ecology* 3: 3–4, 163–173.

Miller, H. & S.-L. Shaw (2001). *Geographic information systems for transportation, principles and applications.* 458 p. Oxford University Press, Oxford.

Mitchell, C. (2004). Making sense of counterurbanization. *Journal of Rural Studies* 20: 1, 15–34.

MRA (1999). Maankäyttö- ja rakennusasetus 10.9.1999/895. Available at <http://www.finlex.fi/fi/laki/ajantasa/1999/19990895> Accessed August 2012

MRL (1999). Maankäyttö- ja rakennuslaki 5.2.1999/132. Available at <http://www.finlex.fi/fi/laki/ajantasa/1999/19990132> Accessed August 2012

Nechyba, T. & R. Walsh (2004). Urban sprawl. *Journal of Economic Perspectives* 18: 4, 177–200.

Nijkamp, P. (1986). Infrastructure and regional development: a multidimensional policy analysis. *Empirical Economics* 11: 1, 1–21.

OECD Territorial Reviews: Finland (2005). OECD (The Organisation for Economic Co-operation and Development) 216 p. OECD, Paris.

Patacchini, E. & Y. Zenou (2009). Urban sprawl in Europe. *In* Burtless, G. & J. Rothenber (eds.): *Brookings-Wharton Papers on Urban Affairs 2009,* 125–149. The Brookings Institution, Washington DC.

Pijanowski, B., D. Brown, B. Shellito & G. Manik (2002). Using neural networks and GIS to forecast land use changes: a Land Transformation Model. *Computers, Environment and Urban Systems* 26: 6 553–575.

Pitkänen, K. (1988) Väestöntutkimus ja yhteiskunta. *Suomen väestötieteen yhdistyksen julkaisuja* 11. 223 p

Pitkänen, K. (2007). Suomen väestön historialliset kehityslinjat. *In* Koskinen, S., T. Martelin, I.-L. Notkola, V. Notkola, K. Pitkänen, M. Jalovaara, E. Mäenpää, A. Ruokolainen, M. Ryynänen & I. Söderling (eds): *Suomen väestö,* 41–75. Gaudeamus, Helsinki University Press, Helsinki.

Population by industry (1979). Central Statistical Office of Finland. Statistical Surveys. N:o 63. Central Statistical Office of Finland (Statistic Finland). Helsinki.

Population projection (2012). Statistic Finland. Helsinki, Available at <http://www.stat.fi/meta/til/vaenn_en.html> Accessed August 2012

Python Programming Language (2012). Python Software Foundation. Available at <http://www.python.org/> Accessed August 2012

Quinet, E. & R. Vickerman (2004). *Principles of Transport Economics.* 385 p. Edward Elgar Publishing, Cheltenham.

Railway statistics (1989). VR-Yhtymä Oy, Helsinki.

Railway statistics (1990). VR-Yhtymä Oy, Helsinki.

Rappaport, J. (2008). Consumption amenities and city population density. *Regional Science and Urban Economics* 38: 6, 533–552.

Rappaport, J. (2009). The increasing importance of quality of life. *Journal of Economic Geography* 9: 6, 779–804.

Rasila, V. (1982). Liikenne. *In* Ahvenainen, J., E. Pihkala & V. Rasila (eds.): *Suomen taloushistoria 2*, 132–153. Kustannusosakeyhtiö Tammi, Helsinki.

Rietveld, P. & P. Nijkamp (1992). Transport and Regional Development. *Serie Research Memoranda* 1992-50. 24 p.

Rietveld, P. & P. Nijkamp (1993). Transport and regional development. In Polak, J. & A. Heertje (eds.): *European transport economics.*, 130–151. Blackwell, Oxford.

Rietveld, P. & F. Bruinsma (1998). *Is Transport Infrastructure Effective?* 383 p. Springer, Berlin.

Rusanen, J., T. Muilu, A. Colpaert & A. Naukkarinen (2003). Georeferenced data as a tool for monitoring the concentration of population in Finland in 1970–1998. *Fennia* 181: 2, 129–144.

Schmitt P., A. Dubois, J. Roto, J. Sterling & C. Schürmann (2008). Exploring the Baltic Sea Region – On territorial capital and spatial integration. *Nordregio Report* 2008:3. 138p .

Seppinen, I. (1992). *Valtaväylä Suomeen, Liikenneministeriö 100 vuotta.* 300 p. VAPK-kustannus and Liikenneministeriö, Helsinki.

Song, S. (1996). Some tests of alternative accessibility measures: a population density approach. *Land Economics* 72: 4, 474–482.

Song, Y., K. Lee, W. Anderson & T. Lakshmanan (2012). Industrial agglomeration and transport accessibility in metropolitan Seoul.*Journal of Geographical Systems* 14: 3, 299–318.

Spiekermann, K. & J. Neubauer (2002). European Accessibility and Peripherality. *Nordregio Working Paper* 2002:9. 43 p.

Spiekermann, K. & H. Aalbu (2004). Nordic Peripherality in Europe. *Nordregio Working Paper* 2004:2. 30 p.

Spiekermann, K. & M. Wegener (2006). Accessibility and Spatial Development in Europe. *Scienze Regionali* 5: 2, 15−46.
Spiekermann, K. & M. Wegener (2007). Update of Selected Potential Accessibility Indicators. *ESPON, Final Report* February 2007. 29 p.
Spiekermann, K., M. Wegener, V. Květoň, M. Marada, C. Schürmann, O. Biosca, A. Segul, H. Antikainen, O. Kotavaara, J. Rusanen, D. Bielańska, F. Fermi, D. Fiorello, T. Komornicki & P. Rosik (2011). TRACC, Transport Accessibility at Regional/Local Scale and Patterns in Europe. *Interim Report* 21/02/2011. ESPON. 164 p. Available at <http://www.espon.eu/export/sites/default/Documents/Projects/AppliedResearch/TRACC/TRACC_Interim_Report_210211.pdf> Accessed August 2012
Suomen teiden historia (1974). Tie- ja vesirakennushallitus ja Suomen tieyhdistys, Helsinki.
Tarkistetut valtakunnalliset alueidenkäyttötavoitteet (2009). Available at <http://www.ymparisto.fi/download.asp?contentid=98972&lan=fi> Accessed August 2012
Taylor, B. (2004). The politics of congestion mitigation, *Transport Policy* 11: 3, 299–302.
TeleFOT (2012). Available at <http://www.telefot.eu/oulufot/> Accessed August 2012
Tervo, H. (2005). Regional policy lessons from Finland. *In* Felsenstein, D. & B. A. Portnov (eds.): *Regional Disparities in Small Countries*, 267–282. Springer, Berlin.
Toivonen, T., T. Jaakkola & M. Vuori (2010). Solmukohta vai pussinperä? *Helsingin yliopiston maantieteen laitoksen julkaisuja B* 55. 62 p.
Transport and communications statistical yearbook for Finland (2010). Statistics Finland. Statistic Finland, Helsinki.
Transport policy guidelines and transport network investment and financing programme until 2020 (2008). Ministerial Working Group on Transport and Communications. Government transport policy report to Parliament. Publications of The Ministry of Transport and Communications 30/2008 Available at <http://www.lvm.fi/fileserver/3008.pdf> Accessed August 2012
Tykkyläinen, M. (1981). Accessibility in the provinces of Finland. *Fennia* 159: 2, 361–396.
Työlliset toimialan (TOL2008) mukaan alueittain 2007-2009 (2012). Statistics Finland. Available at <http://pxweb2.stat.fi/database/StatFin/vrm/tyokay/tyokay_fi.asp> Accessed August 2012
Uimonen, S. (2010). Measuring the highway capital in Finland 1900–2009. *Tampere Economic Working Papers Net Series* 81. 86 p.
Vahtola, J. (2003). *Suomen historia.* 496 p. Otava, Helsinki.
Valtioneuvoston päätös valtakunnallisista alueidenkäyttötavoitteista (2000). Available at <http://www.ymparisto.fi/download.asp?contentid=94382&lan=fi> Accessed August 2012
Valtionrautatiet 1962–1987 (1987). Rautatiehallitus, Helsinki.
Vartiainen, P. (1989). Yhteiskunnan muutos ja yhdyskuntarakenne. *Occasional papers, University of Joensuu, Human Geography and Planning* 13. 111 p.
Vaturi, A., B. Portnov & Y. Gradus (2011). Train access and financial performance of local authorities: greater Tel Aviv as a case study. *Journal of Transport Geography* 19: 2, 224–234.
Verburg, P., G. Koning, K. Kok, A. Veldkamp & J. Bouma (1999). A spatial explicit allocation procedure for modelling the pattern of land use change based upon actual land use. *Ecological Modelling* 116: 1, 45–61.
Verburg, P., J. Ritsema van Eck, T. Nijs, M. Dijst & P. Schot (2004). Determinants of land-use change patterns in the Netherlands. *Environment and Planning B: Planning and Design* 31: 1, 125–150.
Verbyla, D. (2002). *Practical GIS Analysis.* 294 p. Taylor & Francis, London.
Vickerman, R., K. Spiekermann & M. Wegener (1999). Accessibility and economic development in Europe. *Regional Studies* 33: 1, 1–15.
von Thünen J. H. (1826). *Der Isolierte Staat in Beziehung auf Landschaft und Nationalökonomie.* (Translated by Wartenberg C.M. (1966). *Von Thünen's Isolated State.* Oxford: Pergamon Press.
Voss, P. (2007). Demography as a spatial social science. *Population Research and Policy Review* 26: 5–6, 457–476.
Vuoristo, K.-V. (1967). Suomen maaliikenneverkkojen analysointia. *Terra* 79: 3, 87–94.
Watson, M. (1978). The Scale Problem in Human Geography. *Geografiska Annaler B 60*: 1, 36–47.
Weber, A. (1909). *Urber don Standort der Industrien.* MIT Press. Cabridge.
Wegener, M. & D. Bökemann (1998). The SASI model. *SASI deliverable D8 report to the European Commission.* 58 p.
Wegener, M. (2004). Overview of land-use transport models. *In* Hensher, D. A. & K. Button (eds.): *Transport geography and spatial Systems*, 127–146. Pergamon/Elsevier Science, Kidlington.

Wegener, M. & F. Fürst (2004). Land-Use Transport Interaction: State of the Art. *Berichte aus dem Institut für Raumplanung* 46. 119 p.
Wood, S. (2006). *Generalized Additive Models.* 391 p. Chapman & Hall/CRC, London.
Yearbook of Finnish transport statistics (1958). Publications of the transport council. Helsinki.
Yearbook of Finnish transport statistics (1960). Publications of the transport council. Helsinki.
Yearbook of transport statistics (1971). Official statistics of Finland. Publications of the traffic council. Helsinki.
Yearbook of transport statistics (1980). Central Statistical Office of Finland. Central Statistical Office of Finland, Helsinki.
Yearbook of transport statistics for Finland (1990). *Central statistical office of Finland Statistics.* Helsinki.
Yee, T. & N. Mitchell (1991). Generalised additive models in plant ecology. *Journal of Vegetation Science* 2: 5, 587–602.
Youshida, N. & U. Deichmann (2009). Measurement of accessibility and its applications. *Journal of Infrastructure Development.* 1: 1, 1–16.
Zetterberg, S. (2011). *Yhteisellä matkalla.* 543p. Werner Söderström osakeyhtiö, Helsinki.