KU LEUVEN
FACULTY OF SCIENCE

VRIJE UNIVERSITEIT BRUSSEL
FACULTY OF SCIENCE AND
BIO-ENGINEERING SCIENCES

Analysis and modelling of population evolutions in rural areas

Case Study: rural population dynamics and rural abandonment in the Marne Department

Bram VANDENINDEN

Promotor: Prof. A. Van Rompaey
KU Leuven

Thesis presented in
Fulfillment of the requirements
For the degree of Master of Science
in Geography

Academic Year 2016-2017

List of Abbreviations

ABC Agent Based model -Calibrated

ABI Agent Based model - Initial

ABM Agent Based Model

ABV Agent Based model - Validated

AMR Annual migration rate

ASL Above Sea Level

DEM Digital Elevation Model

DEP Delta Employment Potential

DSL Delta Service Level

DTS Distance To the (nearest) Supermarket (from the center of the village)

DTC Distance To the (nearest) City (from the center of the village)

DTH Distance To the (nearest) High school (from the center of the village)

EP Employment Potential

Eq. Equation

GIS Geographic Information Systems

HP Housing Prices

NSE Nash-Sutcliffe Efficiency

PM Particular Matter

RMSE Root Mean Square Error

SL Service Level

SRM Stepwise Regression model

UL Utility Level

UTS Utility To Stay

Yr. Year

Preface

Years ago, I opted for geography because it is a broad program including various important and interesting aspects of the world and science. Looking at certain problems in the world today (e.g. climate change or rural abandonment), you need often a combined knowledge of the disciplines (e.g. GIS, weather and climate, numerical modelling, statistics, computer language, economy, ecology, sociology, landscape analysis, policy, demographics, terrain analysis, …) that were included in the geography program. And that is what I find one of the most interesting aspects of geography and what to my opinion shows the usefulness and importance of geography.

As a conclusion to the geography program, I have written this master thesis related to population evolutions in rural areas in Europe. I have conducted this research with a lot of enthusiasm and a strong determination. I have tried to never forget to look at existing researches as to my own research with sufficient critical reflection. I am looking forward to graduate as Master of Science in Geography and to offer my knowledge on the job market in order to contribute to progress in science and society.

I want also to thank some people that have contributed to the end result of this research. First, my gratitude goes to my promotor, professor Van Rompaey, for providing this thesis topic, his assistance during the first two days of my field work and his useful advice during the research process of writing this thesis. Further, I want also thank my colleagues and friends: Laurens, for his often useful advice from time to time and everyone for their presence at places such as the library. Because of them, writing the thesis was not a lonely process in desolation, but were there also moments of fun and possibilities to share frustrations. Next, I want also to thank my parents and family.

Abstract (English)

Over the past few decades, rural depopulation has become a common phenomenon that occurs in many places across the globe. Especially in the United States, Europe and China there are many regions experiencing rural depopulation. In some cases the depopulation can evolve that far that certain rural villages, especially remote rural villages located far from regional cities, become completely or almost completely abandoned. This depopulation has many consequences leading to a reduced quality of life for the remaining people. The consequences with the greatest impact are a lack of employment (or long commuting distance to employment place), social exclusion, low-quality education and the disappearing of services in those rural villages. This includes services such as schools, mini markets and bakeries, and also services like recreational activities and public transport.

To get a more profound insight in population dynamics in rural areas, this research provides an in-depth analysis of population evolutions in rural areas on the basis of a case -study. The selected case is the Marne department in France. This area is suitable for the research because there are high dynamics within this department: some rural regions and rural villages within the Marne department experience strong depopulation while other rural regions and rural villages in the department are characterized by an increasing population. The eastern and southern regions of the department experience strong depopulation.

For the research, large amounts of data were gathered and analyzed. During several weeks of field work, data regarding the characteristics of the population and the characteristics of the villages were gathered. Additionally, many data from various existing databases are used. When it comes to the characteristics of rural villages, the main triggering factors (push factors) for young and middle-aged people to migrate are mainly the employment potential and to a lesser extent the service level of the villages. For elderly retired people, the services level in a village becomes the most important triggering factor for migration. The main factors explaining the attractiveness of rural villages (pull factors) once people have decided to migrate are the employment potential, the service level of a village and the distance to the nearest supermarket from the village and the HP in a rural village. Some feedback mechanisms are clearly present in rural population dynamics: in depopulating villages, many services start closing once the population falls below a threshold of around 2500 to 300 people, which causes a positive feedback on the depopulation, the services closing makes the villages less attractive, which accelerates depopulation. Housing Prices are a decelerating factor: they are low in villages that are depopulating and where services tend to close. But this has a decelerating effect on depopulating, seeing as poor people who cannot afford to buy a house in more attractive areas, are in some cases forced to turn their eye towards houses in depopulating rural villages.

Related to the characteristics of the population, there is found a relationship between age and education and the residential mobility: younger-and high educated people are more eager to migrate. Also there is found a significant relation between the migration type and the education level: highly educated people choose predominantly rural-urban migrations, while in the average educated people most often choose for rural-rural migrations and poorly educated people choose either rural-rural or urban-rural migrations.

A second part of the study takes a closer look at whether it is possible to simulate rural population dynamics with an agent-based model. An initial model was developed and calibrated based on observed data. There is found an agent-based model can be very useful in simulating rural population evolutions. The model performs significantly superior compared to a null model, a chance model, a model using the observed mean of the simulation period as prediction and the agent-based model also performs significantly better compared to a stepwise regression model. This implies agent-based modelling can definitely have a place in population modelling. However, improvements are needed to get a very accurate simulation of the population. Possible suggestions for improvements are more extensive and longer field work periods to gather information about the education level of at least a few thousand respondents. Also combining agent-based modelling and big data analysis probably can the model cause performing more accurate.

The agent-based model predicts that rural depopulation in the Marne department will get worse, with the majority of rural villages experiencing moderate to strong depopulation by 2040. In the east of the department, where at present the majority of the rural villages already struggles with depopulation, the situation is expected to worsen. Nearly every village in this region is expected to experience depopulation by 2040. The eastern part of the Marne department is the region where the characteristics of the villages are the least attractive: almost no employment opportunities, a low number of services, a great distance to the closest supermarket and the absence of public transport.

In some areas, examples show that it is possible to turn back the process of rural depopulation. Sometimes through spontaneous repopulation, but more often through controlled repopulation. Converting a fraction of the rural villages into either an attractive destination for tourists or into 'ecovillages' can be successful. However, there should be debate whether it remains sustainable to invest equally in all rural villages or if it may be preferable to let some rural villages collapse under controlled circumstances.

Abstract (Nederlands)

Ontvolking in rurale gebieden is in de laatste decennia een veel voorkomend verschijnsel in verschillende regio's in de wereld. Zeker in de Verenigde Staten, Europa en China zijn er verschillende regio's die te kampen hebben met rurale ontvolking. In sommige gevallen kan de ontvolking zover evolueren dat sommige rurale dorpen, zeker de afgelegen dorpen, volledig of bijna volledig verlaten worden. Voor de achtergebleven inwoners kan dit leiden tot een gereduceerde kwaliteit van leven. De meest merkbare gevolgen zijn een gebrek aan tewerkstelling (of een lange pendeltijd tot de plaats van tewerkstelling), sociale exclusie, lage kwaliteit van onderwijs en het verdwijnen van diensten uit deze rurale dorpen. Dit zijn diensten zoals scholen, kleine winkels, bakkers, en ook diensten zoals ontspanningsactiviteiten en het openbaar vervoer. Om een dieper inzicht in de dynamieken van rurale ontvolking te verkrijgen, wordt er in dit onderzoek een diepteanalyse uitgevoerd van de populatie-evolutie in rurale gebieden op basis van een gevalsstudie. De geselecteerde case in dit onderzoek is het Marne departement in Frankrijk. Dit gebied is geschikt omdat er grote dynamieken zijn binnen het departement: sommige rurale gebieden en rurale dorpen in het departement zijn gekenmerkt door een sterke ontvolking, in andere regio's van het departement is er dan weer een stijgende populatie. Regio's met sterke ontvolking zijn gelokaliseerd in het oosten en zuiden van het Marne departement.

Gedurende verschillende weken van terreinwerk zijn er data omtrent de karakteristieken van de bevolking en karakteristieken van de dorpen verkregen. Bovendien werden ook grote hoeveelheden data van verschillende bestaande databases gebruikt. De belangrijkste factoren gerelateerd aan de kenmerken van de gemeente die een trigger voor ontvolking kunnen zijn, zijn voornamelijk het tewerkstellingspotentiaal en in mindere mate het aantal voorzieningen in een dorp. Voor ouderen (gepensioneerden) echter, is het aantal voorzieningen de meest belangrijke factor die emigraties kan uitlokken. Eenmaal mensen hebben beslist om te migreren, zijn de volgende factoren de belangrijkste die de attractiviteit van rurale dorpen verklaren: het tewerkstellingspotentiaal, het aantal voorzieningen en de afstand tot de dichtstbijzijnde supermarkt en de huizenprijzen in het dorp. Sommige feedbackmechanismen zijn duidelijk aanwezig in het proces van rurale ontvolking: in dorpen die leeglopen, beginnen veel voorzieningen te sluiten eenmaal het aantal mensen in het dorp onder de grens van 300 komt. Door het sluiten van die voorzieningen wordt het dorp minder aantrekkelijk en zal de ontvolking er versnellen. De huizenprijzen zijn laag in dorpen die leeglopen en waar er weinig of geen diensten meer zijn. Dit vertraagt de ontvolking: in sommige gevallen migreren armere mensen, die geen huis in meer aantrekkelijke plaatsen kunnen betalen, naar dorpen waar de huizenprijzen laag zijn.

Gerelateerd aan de kenmerken van de bevolking is er een verband gevonden tussen de leeftijd, het opleidingsniveau en de mobiliteit van de bevolking: jongere mensen en hoogopgeleide mensen migreren vaker. Ook is er een significant verband tussen het type migratie dat mensen kiezen en hun opleidingsniveau: hoogopgeleide mensen kiezen vaak voor migraties van het platteland naar de stad, gemiddeld opgeleide mensen kiezen vaak voor migraties van het platteland naar het platteland en laag opgeleide mensen kiezen voor ofwel migraties van het platteland naar het platteland ofwel migraties van de stad naar het platteland (als ze initieel in de stad woonden).

In een tweede deel van dit onderzoek is er onderzocht of het mogelijk is om rurale populatiedynamieken te simuleren met een agent-based model. Zo een model kan zeer nuttig zijn in het simuleren van populatie-evoluties. Het model scoort significant beter vergeleken met een nul-model, een kansmodel, een model dat het geobserveerd gemiddelde van de simulatieperiode als voorspelling gebruikt en het scoort ook beter dan een stapsgewijs regressie model. Dit betekent dat agent-gebaseerde simulaties zeker een rol hebben in het modelleren van populatie-evoluties. Echter, verbeteringen zijn nodig om een zeer accurate simulatie van de bevolking te bekomen. Mogelijke suggesties voor verbeteringen zijn een uitgebreidere en langere terreinwerkcampagne om meer informatie over het opleidingsniveau van minstens enkele duizenden mensen te bekomen. De simulaties van de populatie kunnen zeer waarschijnlijk meer accuraat worden wanneer agent-based modelleren en een analyse van ‘big data’ gecombineerd worden.

Het agent-gebaseerde model voorspelt een verergering van de ontvolking voor het Marne departement waarbij in 2040 meer dan de helft van de rurale dorpen een matige tot sterke ontvolking zou ondervinden. In het oosten van het departement, waar op heden reeds een meerderheid van de rurale dorpen kampt met ontvolking, wordt voorspeld dat de situatie het meest ernstig zal zijn. Bijna alle rurale dorpen in dit deel van het departement zullen ontvolking kennen in 2040. Het oosten van het departement komt ook overeen met de regio van het Marne departement waar de kenmerken van het dorp het minst attractief zijn: een laag aantal tewerkstellingsmogelijkheden, weinig voorzieningen, een grote afstand tot de supermarkt en afwezigheid van openbaar vervoer. Er zijn voorbeelden die illustreren dat het mogelijk is om het proces van rurale ontvolking om te keren op sommige plaatsen. In sommige gevallen door natuurlijke aangroei van de populatie, maar vaker door gecontroleerde populatie-aangroei. De conversie van een fractie van de rurale dorpen in attractieve toeristenbestemmingen of de conversie van rurale dorpen naar ‘ecodorpen’ kan succesvol zijn. Echter, er is debat nodig of het duurzaam blijft om in alle rurale dorpen tegelijk te investeren of dat het beter is om een aantal rurale dopen te laten imploderen onder gecontroleerde toestand.

Abstract (Français)

La dépopulation dans les zones rurales est un phénomène devenu commun au cours des dernières décennies. Cette tendance s'observe aux différent parts du globe, notamment dans de nombreuses régions des États-Unis, d'Europe et de Chine. Dans certains cas, l'évolution de la dépopulation est telle que certains villages ruraux, et plus particulièrement les villages isolés, deviennent en partie voire totalement abandonnés. Cette situation engendre de nombreuses conséquences et peut mener à la diminution de la qualité de vie des habitants qui restent. Les répercussions les plus perceptibles sont la pénurie d'emplois (ou de longs trajets entre le domicile et le lieu de travail), l'exclusion sociale, la moindre qualité de l'enseignement et la disparition des services tels que les écoles, les petits magasins, les boulangeries, les activités de détente et les transports en commun dans ces villages ruraux.

Afin de mieux comprendre les dynamiques de la dépopulation rurale, une analyse approfondie de l'évolution de la population dans les zones rurales sera réalisée dans le cadre de la présente recherche, sur la base d'une étude de cas. Le cas sélectionné est le département de la Marne, en France. Cette région convient bien pour la recherche car on y observe d'importantes dynamiques. En effet, certains villages et régions de ce département sont caractérisés par une forte tendance à la dépopulation tandis que d'autres enregistrent une croissance de leur population. Les régions marquées par une dépopulation notable sont l'est et le sud du département de la Marne.

Pour la recherche, de grandes quantités de données ont été recueillies et analysées. Pendant plusieurs semaines de travail sur le terrain, des données relatives aux caractéristiques de la population et des villages ont été rassemblées. En outre, de nombreuses informations issues de sources internet ont également été utilisées. Les principaux facteurs liés aux caractéristiques des villages pouvant déclencher une dépopulation sont le potentiel d'emploi et, dans une moindre mesure, le nombre de structures présentes dans le village. Toutefois, pour les personnes âgées (à la retraite), la présence de structures constitue le facteur le plus important pouvant entraîner une émigration. Une fois que les habitants ont décidé de s'exiler, les principaux facteurs suivants peuvent expliquer l'attractivité des villages ruraux : le potentiel d'emploi, le nombre de structures, la distance par rapport au supermarché le plus proche et le prix des logements dans le village. Plusieurs mécanismes de rétroaction ont clairement un rôle à jouer dans le processus de dépopulation rurale : dans les villages en cours de dépeuplement, de nombreuses structures commencent à fermer leurs portes une fois que le nombre d'habitants descend sous la barre des 300. Cette situation entraîne une attractivité moindre du village et une accélération de la dépopulation. Néanmoins, le prix des logements permet de ralentir ce phénomène. En effet, les prix diminuent dans les villages désertés où peu voire

pas de structures subsistent. Cette chute des prix freine la dépopulation car dans certains cas, les citoyens pauvres qui ne peuvent pas se permettre d'acheter une maison dans des régions plus attrayantes s'exilent dans des villages où les prix des logements sont à la baisse.

Une deuxième partie de la présente étude analysera s'il est possible de simuler les dynamiques de la population rurale par le biais d'un modèle multi-agents. Il apparaît que ce type de modèle peut s'avérer très utile dans le cadre de la simulation des évolutions de la population et donne de meilleurs résultats qu'un modèle nul, un modèle aléatoire, un modèle que la moyenne observée au cours de la période de simulation utilise comme prédiction ou un modèle de régression séquentielle. Cela suppose que les simulations multi-agents jouent clairement un rôle dans la modélisation des évolutions de la population. Toutefois, des améliorations sont nécessaires afin de pouvoir obtenir une simulation très précise de la population. Parmi les éventuelles suggestions d'amélioration figurent une campagne de travail sur le terrain plus vaste et plus longue afin de recevoir davantage d'informations sur le niveau de formation de plusieurs milliers de personnes au minimum. Les simulations de la population peuvent très probablement être plus précises si la modélisation multi-agents peut être combinée à l'analyse de gros volumes de données.

Le modèle multi-agents prévoit une aggravation de la dépopulation dans le département de la Marne. En effet, il est estimé qu'en 2040, plus de la moitié des villages ruraux pourraient subir une dépopulation modérée à considérable. L'on prévoit que c'est dans l'est du département, où une majorité des villages ruraux luttent déjà actuellement contre la dépopulation, que la situation sera la plus grave. La quasi-totalité des villages ruraux situés dans cette région du département seront touchés par la dépopulation d'ici 2040. L'est du département correspond également à la région de la Marne où les caractéristiques des villages sont les moins attractives : peu de perspectives d'emploi, peu de structures, une longue distance jusqu'au supermarché le plus proche et une absence de transports en commun.

Dans certaines régions, des exemples montrent qu'il est possible d'inverser le processus de dépopulation rurale, parfois par un accroissement naturel de la population, mais le plus souvent par un repeuplement contrôlé. La conversion de certains villages ruraux en destinations touristiques attrayantes ou en « éco-villages » peut être une solution efficace. Néanmoins, un débat s'avère nécessaire pour déterminer s'il demeure viable d'investir de manière égale dans l'ensemble des villages ruraux ou s'il serait plus judicieux de laisser certains villages imploser dans un environnement contrôlé.

Table of contents

List of figures

List of tables

CHAPTER 1 : INTRODUCTION AND PROBLEM STATEMENT

1.1. Recent population dynamics in the global north

During centuries, settling in rural areas has been the dominant way of living in western societies. Before the Industrial revolution was initiated, some cities existed, and around 1800 circa 20% of the European population lived in urban areas. In 1950 about 50% of the European population lived in cities (Lampard, 2016). In the same time span the global urban population raised from 2% to 13%. This transition was due to mechanical improvements leading to increased food production levels. As consequence, it was no longer necessary for the whole population to be involved in agriculture. Figure 1 illustrates the evolution of the share of the rural and urban population in Europe and the world. Within Europe, there are strong variations in urbanization trends (figure 2): northern and western regions of Europe were characterized by an earlier urbanization while southern and eastern regions of Europe have been catching up during the last decades. Data on urbanization number should be interpreted with care since there is no clear definition on what an urban zone or an urban population is. Some definitions are based on the absolute population number of a settlements or on the degree of spatial clustering of houses, but in the literature various thresholds are used for different regions and countries (Satterthwaite, 2007).

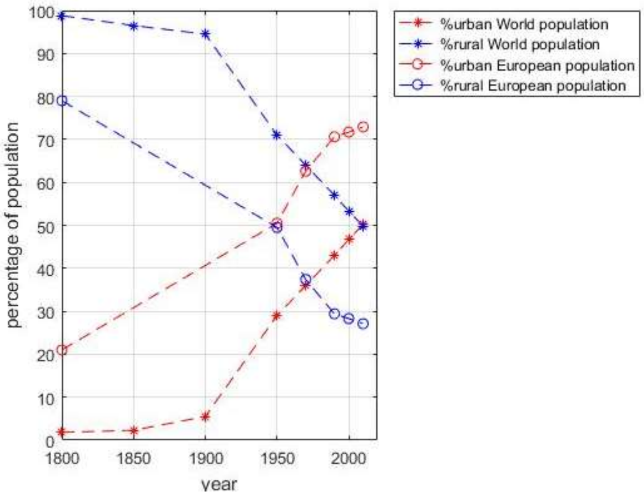


Figure 1: Share of urban population for Europe and the world over time. Own processing based on data from combining various data sources (Lampard, 2016; Satterthwaite, 2007).

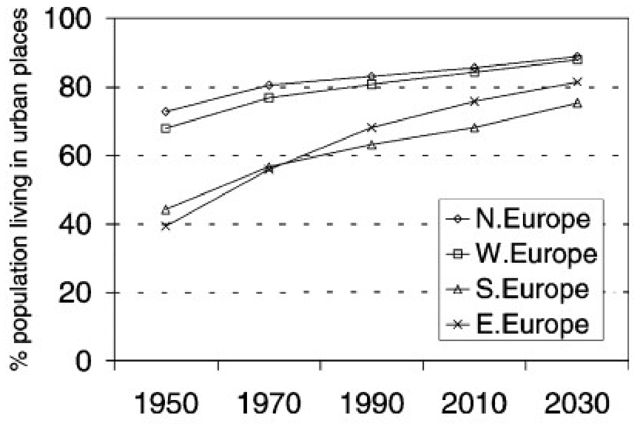


Figure 2: Urbanization level over time in different parts of Europe (Antrop, 2004). For the data point of 2030, projections from the United Nations are used.

The process of urbanization has strong effects on the rural villages. Villages in the proximity of urban zones or villages with a good connection to the urban zone are mostly characterized by a strong population growth, while remote villages often lose population (Antrop, 2004). The process of population increase in the hinterland of an urban center is called suburbanization.

After the Second World War, suburbanization appeared as a new phenomenon in Europe. The characteristic of suburbanization is that the population decreases in the city center and increases in suburban areas. An example of this is the city of Paris: the inner city was region with the highest increase in population in the 1950's while from the 1970's onwards it was the region in France losing the most population (Antrop, 2004; Baccaïni, 2007). This process of suburbanization also affects rural villages: often rural villages located in places near the city grow strongly while a fraction of more remote rural villages will depopulate.

During the last decade, the population in Europe is still growing, but the increase in population is decelerating. The net population evolution at a certain place equals the sum of the natural growth of the population and the migrations (equation 1). In 2015, the population increase is estimated at 3,5%, with a negative natural growth of the population of -0,3% and a net migration of +3,7% (Eurostat, 2016).

$$net\ popopulation\ change = natural\ growth + net\ migration\ balance \qquad \text{(eq. 1)}$$

As shown on the map of Europe (figure 3), the population growth is not equal in all European countries. There are countries with a strong population growth while other countries are characterized by strong population losses. The cause of the net population differs between countries: while some countries such as France gain population because of natural growth related to a higher fertility rate than the European average, other countries such as Germany are characterized by strong negative natural population growth. In Germany, the net population evolution is net positive because of a very high number of migrations to the country, especially towards the surroundings of Berlin and München (figure 3 : net population evolution, figure 4: natural growth of the population, figure 5: net migration balance).

A spatial aspect related to the physical environment can be detected: the population density is in most cases increasing in coastal areas below a latitude of about 60°N (e.g. France, Spain, Italy, Bulgaria, Slovenia, Croatia, Portugal) . Above this latitude (the Scandinavian countries and some Eastern countries like Latvia and Lithuania) , no increase in population is registered (figure 3).

The European population change maps clearly show the urban-rural divide: countries like Spain and Germany have a net decrease in population in most regions, except in the surroundings of their capitals (Berlin and Madrid) and some other important cities (like München and Barcelona). In most cases the population increase in the (sub)-urban areas outweighs the losses in the rural areas (in Germany the net population evolution is net positive, in Spain the net population evolution is almost constant (Eurostat, 2016)).

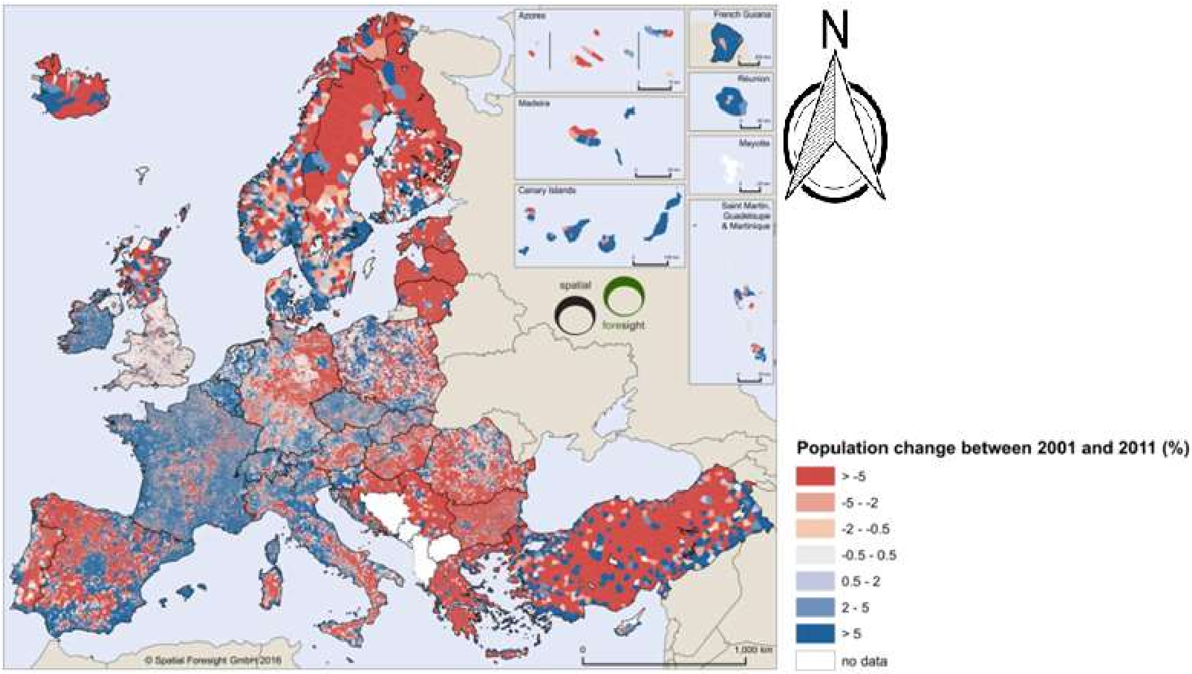


Figure 3: Population change in Europe between 2001 and 2011 (adapted from Spatial Foresight, 2016).

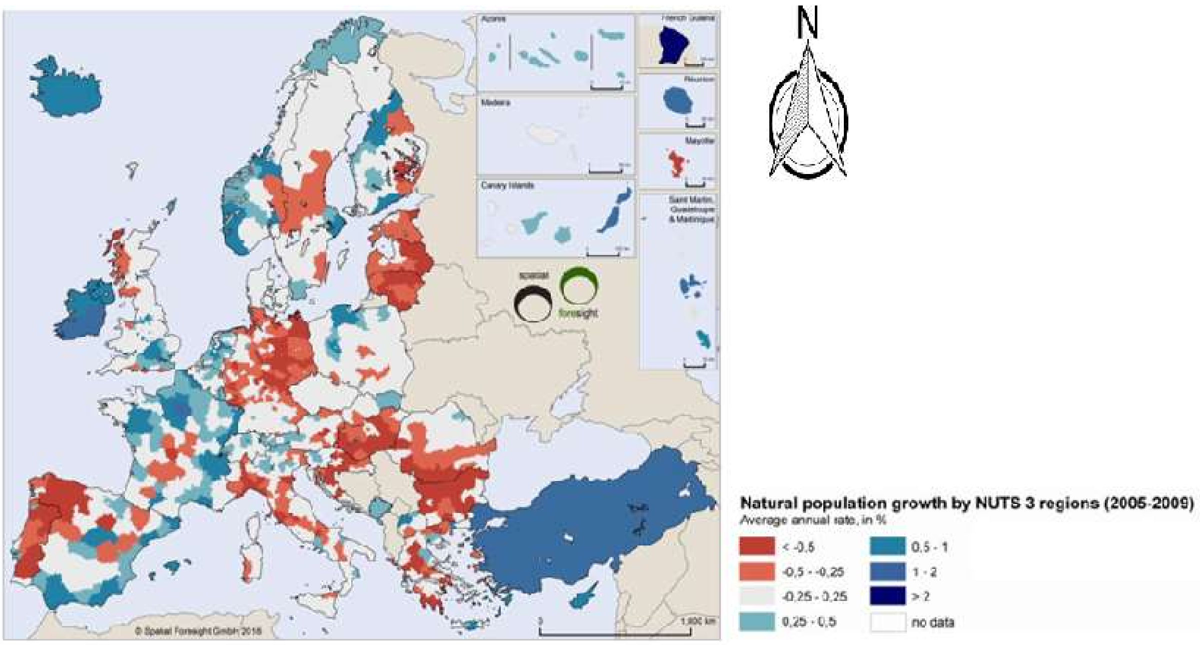


Figure 4: Natural population growth in Europe (adapted from Spatial Foresight, 2016).

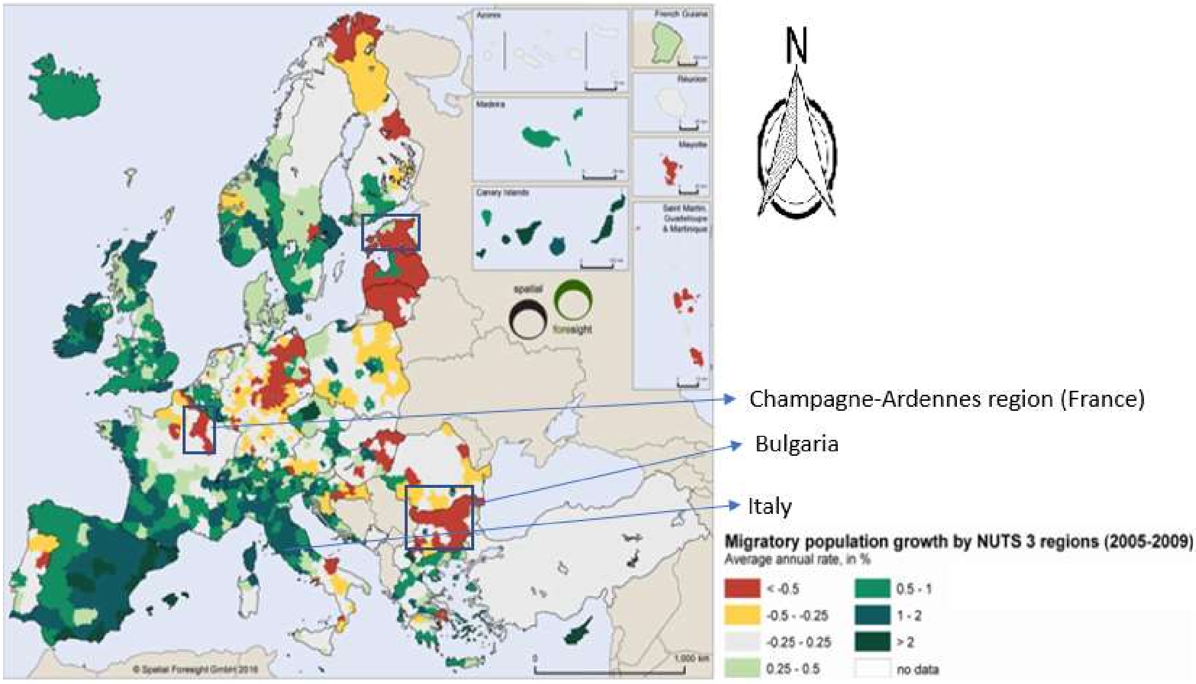


Figure 5: Net migration balances in European countries (adapted from Spatial Foresight, 2016).

Not only between countries the evolution of the population shows an important variability, also within countries, there is a lot of spatial differentiation. In many European countries, a process of rural abandonment can be observed. The rural population is declining in many regions while the urban and suburban population is rising. These trends can be correlated with the growth of economic activities in the (sub)urban areas and the decline of the rural economy in the rural areas. Next, some interesting examples of rural abandonment are highlighted that clarity the phenomenon of rural abandonment.

France, where there is net a moderate increase of the population due to a high fertility rate (causing high natural growth of the population), is one of the European countries where a very strong rural depopulation occurs, especially in the north and the center of France (in the south and in coastal regions, there is a positive migration balance, figure 5). This region is dominantly rural with mainly migrations out of rural areas, however, in some areas, even some cities are starting to decline. (Baccaïni & Levy, 2009) report that between 1999 and 2004 the net number of people leaving the Champagne-Ardennes region was higher than the number of people moving to the region, and this for all age groups. Rural depopulation mainly occurs in villages with a low population density (Liu & Laske, 2010). Measurements taken by the French government to improve the situation in subordinated rural areas are the stimulation of investments in those regions by means of giving tax exemption to companies that are willing to settle in those depopulating rural villages and for professionals who are engaged in activities that are non-commercial, such as physicians and lawyers (Liu & Laske, 2010). The case of France is interesting because of its contrasts: while some rural areas have a relatively high population density with increasing population and a well-developed infrastructure, other rural areas, mostly with a low population density, have a decreasing population and a poor infrastructure and no employment (Liu & Laske, 2010). Also within the depopulating rural areas such as the Champagne-Ardennes region the variability is remarkable: while an important number of the villages has a strong population decrease, other remote rural villages are characterized by a stable or even increasing population number.

In **Bulgaria** (also indicated on figure 5), rural depopulation is prominent present in nearly all areas except the region around the capital city Sofia. However, in the regions with an important rural depopulation, there remain often a small fraction of villages with an increasing population. Both a negative natural growth of the population (figure 4) as net emigrations (figure 5) cause the fall of the population in rural areas (figure 3). Many rural villages have been virtually erased during the last decades. Closing of schools and health care centers is common in rural areas and is expected to worsen in the future. Technical and social infrastructure deteriorated in different places. Bulgarian citizens also lose the right to elect a mayor if the population in a village falls below a certain threshold (Mladenov & Ilieva, 2012). Attempts to slow down the process are mainly focused on the improvement of the transport infrastructure. The purpose of this improvement is to revive daily commuting between urban and rural and to ease the development of various forms of tourism (Mladenov & Ilieva, 2012).

Until present, those measures were not very successful. An example of a country where the rural abandonment trend was reversed in some cases is Italy. Especially between the 1950's until the 1990's, Italy was marked by a period with a lot of rural abandonment with as main cause inter-regional emigrations. From the 90's onwards, the amount of rural abandonment in Italy has decreased due to a decrease of emigrations from those regions. For some rural villages, even revitalization occurred. This revitalization of Italian rural villages can be due to tourism or renewed human activity. Renewed human activity can be either spontaneous as planned trough the development of 'ecovillages' (Di Figlia, 2016). 'Ecovillages' are an example of government-intervened rural vitalization whereby eco-sustainability is the core characteristic of the ecovillage (Di Figlia, 2016).

Rural abandonment, is, however, not limited to Europe. The process of rural abandonment also exists in the Great Plains in **America**, and is even much stronger than in Europe. The great plains are an inland area in the United States with highly postwar inter-regional population dynamics (figure 6). The Great plains have a low population density consisting of 18% of the landmass of the United States and containing only 3% of the population of the united states (Wilson & U.S. Census Bureau., 2009). Even though the population in the Great Plains doubled from 1950 to 2007, with 4.9 million inhabitants in 1950 and 9.9 million inhabitants in 2007, more than half of the counties within the Great Plains are characterized by a decrease of population (figure 6), implying strong population growth in the cities and serious declining population in the most rural counties. In a small fraction of the rural counties the population remains stable or increases (Wilson & U.S. Census Bureau., 2009). The doubling of the population also suggests a stronger population growth compared to European countries. The percentage of people living in cities in the Great Plains increased from 39% in 1950 to 68% in 2007. The population living in pure rural countries decreased from 39% in 1950 to 16% in 2007. (Wilson & U.S. Census Bureau., 2009). One of the main reasons for rural depopulation in the Great Plains is the improvement in farming technology, switching to more efficient farming techniques when old farmers retire. Also the lack of employment in the rural areas in the Great Plains has a role of importance (Balar, 2012). Possible sources of revitalization of a part of the rural counties can be the building of wind farms, which would create new jobs in these rural areas (Balar, 2012).

Also in **China,** rural abandonment is becoming problematic in many places. Urbanization, along with the decline of agricultural employment cause strong amounts of rural depopulation in many areas with negative consequences for the people left in rural areas: no education facilities for children, no cultural or entertainment facilities and for some areas also a lack of security (Li, 2015).

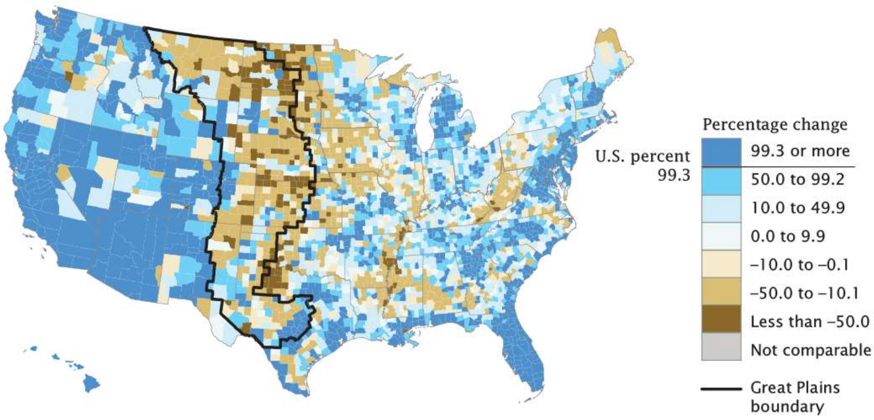


Figure 6: Percentage change in population by County from 1950 to 2007 (Wilson & U.S. Census Bureau., 2009).

From the discussed examples it became clear rural abandonment often has undesired consequences impacting the quality of life of the people. After those specific examples, what follows is a more general description of rural abandonment and its consequences in Europe and existing policies in Europe related to rural abandonment and rural development.

1.2. Rural Abandonment

In some rural villages, rural abandonment can, once initiated, continue for a long time until a village become nearly completely abandonment and eventually becomes a ghost village. If through the process of depopulation the number of people that remain falls below a certain threshold, this has a lot of undesired consequences for the people that stay, possibly seriously impacting their quality of life.

Farmland abandonment and decrease in agricultural employment along with decrease in local industrial employment through des-industrialization are the main direct causes of rural depopulation. A general characteristic of most villages with a declining population, is that they are remote and have decreasing local employment in agriculture and industry. Additionally, public transport is often shut down in those regions, possibly accelerating depopulation (McArthur et al., 2014). There are four main domains where problems are possible to arise when population decreases in rural areas (European Commission, 2008):

- **Demography:** Ageing of the population and emigration of young and well-educated people (European Commission, 2008).
- **Remoteness:** Shut down of services in villages like ending of health services for elderly, closure of shops, childcare, schools, libraries, postal offices and so on (European Commission, 2008). Related to transport, many rural villages suffer from

limited transport infrastructure or policy decisions to shut down public transport (Dammers & Keiner, 2006; European Commission, 2008; McArthur et al., 2014)

- **Education:** School attendance is in the first years of life in many of European countries significantly lower in rural areas compared to urban areas. Especially children from poor and less-educated people are at risk for poverty and social exclusion if they not attend school when they are a few years old. High poverty and low employment rates are often characteristic for countries with a high share of low-educated people (European Commission, 2008).
- **Labor market:** Poverty and social exclusion occur in the agricultural sector because of low incomes and season-dependent work. For these reasons, working in the agricultural sector is very unattractive for young people (Dammers & Keiner, 2006; European Commission, 2008).

The European Commission and the member states of the European union are aware of these problems and have developed policy plans to revive rural areas. Those plans are developed at various levels. At the European level, there is the Common Agricultural Policy (CAP) and also some member states have additionally their own programs for stimulating the development of rural areas (European Commission, 2013, 2015). About 60 billion Euros are spent trough the CAP policy yearly, representing over 40% of the total EU budget (European Commission, 2013). The distribution of the money spent is shown in figure 7. New direct payments (about 70% of the CAP budget) and rural development (about 20% of the CAP budget) are the two main destinations of the budget. New direct payments consist of various direct payments to farmers. Without subsidies, the wages in the agricultural sector are often lower than in the other sectors, so direct payments compensate for this wage gap. Therefore, many agricultural products, like milk, beef and cereals, are subsidized. Second, because the number of young farmers has fallen sharply the last decades, there are pressing questions about the vitality and future of the agriculture sector in Europe. Therefore, young farmers received additional subsidies in the form of direct payments to help them get started. Third, there are ‘green’ direct payments if farmers apply practices which are beneficial for the environment and the climate, like crop diversification and maintenance of permanent grassland (European Commission, 2015).

Rural development (about 20% of the CAP budget) consists of transferring knowledge and innovation in agriculture and forestry in rural areas, enhancing competitiveness of agriculture trough promoting sustainable and innovative technologies, restoring and enhancing ecosystems who are related to forestry and agriculture, supporting climate resilient resources and aiding economic development in rural areas (European Commission, 2013).

While there may be consensus that stimulating those European regions helps the people employed in the agricultural sector, there are various criticisms of the applied policies. According to the Organization for Economic Co-operation Development (OECD), the focus of the European policy is too much about agriculture and not enough oriented towards the other problems in rural areas. OECD therefore proposes (1) to invest more in a better connectivity between urban and rural, (2) to stimulate more economic investments in rural areas that would enhance the overall productivity and (3) to promote a better international policy dialogue about the issue (OECD, 2014). The European Commission invests significant budgets in agriculture, but little attention goes to problems related to remoteness, education, demography and industry.

Another frequent critique is that the EU agricultural policy causes damaging and devastating effects on the economy in developing countries (Boysen et al., 2014). Because of the subsidized agricultural products in the European Union, some countries are able to dump surplus productions at low prices on the international market, with the consequence that sometimes local farmers in developing countries are not able to sell their products because they cannot compete with highly subsidized EU products (Boysen et al., 2014). Additionally, the agricultural activities remain the main cause for the loss of biodiversity within the European Union (European Commission, 2017).

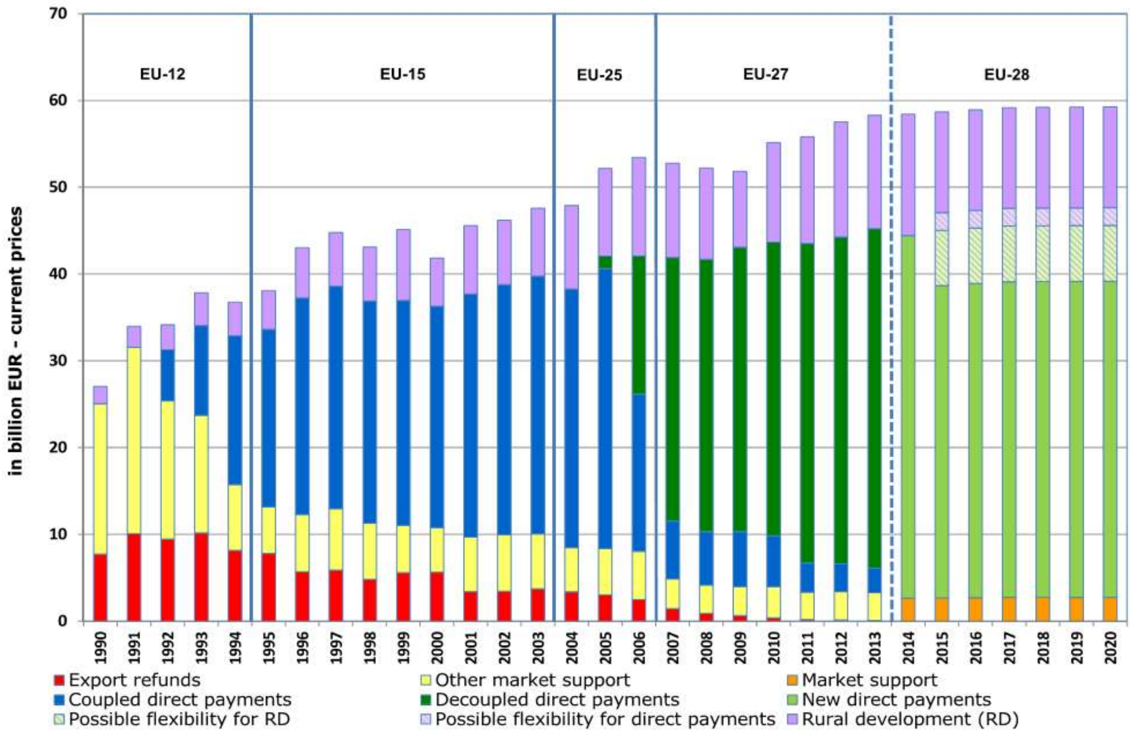


Figure 7: European money spent to support rural areas in Europe under the Common Agricultural Policy (CAP).Evolution of budgets from 1990 to 2020 (European Commission, 2013).

As described, in a lot of regions in Europe, there are rural villages with a decreasing population, but often in those depopulating regions, there remains a (small) fraction of remote rural villages with increasing population. In the next section is examined what the cause are that some rural villages with similar characteristics at the first sight have opposite population evolutions.

1.3. Controlling factors of rural population dynamics

1.3.1. Controlling factors related to the physical environment

Most studies related to population dynamics in Europe do not refer to factors related to the physical environment as possibly influencing factors of population dynamics. However, when properly investigated, a significant relationship[1] was sometimes found. A study conducted in Italy found a relationship between elevation and rural abandonment. Figure 8 shows that mountainous and hilly rural regions are more often abandoned compared to rural regions located on lowland plains. (Di Figlia, 2016).

Disasters such as earthquakes, volcanic eruptions and landslides can also play a pivotal role as controlling factors of population dynamics in some European countries like Italy (Di Figlia, 2016). The presence of water in the form of a sea or ocean also controls population dynamics: rural abandonment rarely occurs along coastal areas in Europe (Baccaïni & Dutreuilh, 2007; Di Figlia, 2016). In the past the climate has influenced European migrations, especially during the little ice age (Pfister & Brázdil, 2006). Certain studies also suggest that future climate change, causing worsening droughts in Mediterranean areas, can accelerate rural abandonment in those regions affected by serious droughts (De Haas, 2011). In studies about population dynamics outside Europe, especially in Africa and Asia, climate is often considered as one of the most important controlling factors of population dynamics (Gray & Mueller, 2012). Table 1 summarizes the factors that were reported to influence population dynamics in Europe.

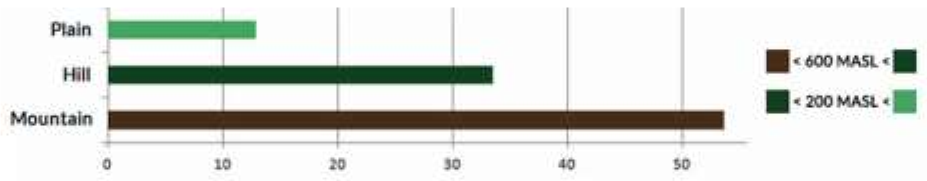


Figure 8: Relative occurrence of abandonment of rural villages in Italy. **Source:** (Di Figlia, 2016).

[1] To avoid confusion, In this research is chosen to only use the word significant when it means there is a statistical significant relationship. In other contexts, the use of the word significant is avoided.

Table 1: *Controlling factors related to the physical environment proven to influence population dynamics in Europe in existing studies*

Controlling factor	Kind of relationship	Place of the study	Source
Elevation above sea level	More depopulation at higher elevation	Italy	(Di Figlia, 2016)
Distance to the sea	Population increase nearby the coast	France, Italy	(Baccaïni & Levy, 2009; Di Figlia, 2016)

1.3.2. Controlling factors related to social and economic situations

In most studies economic factors and especially employment opportunities are considered the most important drivers of population dynamics (De Haas, 2011; McArthur et al., 2014; Zhang & Jager, 2011). The income level of the jobs can also play a role (Zhang & Jager, 2011). Policy interventions can influence those economic factors by investing in regions and stimulating employment creation. Often, transport infrastructure can also play an important role: a Scandinavian study demonstrated, that when in rural regions that depend on industrial employment, when factories closed, population decrease can be reduced or prevented when the road infrastructure is improved before or short after the closing of the factory (McArthur et al., 2014). The presence of a regional city in the surroundings seems to be a factor in some studies with more people tending to migrate towards areas nearby a regional city (Détang-Dessendre et al., 2008). Table 2 summarizes the factors that were reported to influence population dynamics in Europe.

Table 2: *Controlling factors related to socio-economic characteristics proven to influence population dynamics in Europe in existing studies*

Controlling factor	Kind of relationship	Place of the study	Source
Employment opportunities	Employment attracts people	United Kingdom	(Zhang & Jager, 2011).
Income	Higher income in areas with positive population evolution	United Kingdom	(Zhang & Jager, 2011)
Transportation networks and possibilities (roads, public transport, …)	Good road infrastructure and public transport availability prevents depopulation	Norway	(McArthur et al., 2014)
Distance to (large) city, size of closest city	More people tends to migrate towards surroundings of those cities	France	(Détang-Dessendre et al., 2008)
Policy interventions	Can both accelerate as slow down or prevent depopulation	Norway	(McArthur et al., 2014)

1.3.3. Controlling factors related to characteristics of the population and individuals

The likelihood of migration and type of migration is dependent on the individual's age, received education level, type of job and family composition (Détang-Dessendre et al., 2008). Social factors related to the individual also play a role. If individuals have an extensive social network or a small social network in the rural village they live and they consider that network important, they are less likely to migrate (Zhang & Jager, 2011).

Table 3: *Controlling factors related to characteristics of the population proven to influence population dynamics in Europe in existing studies*

Controlling factor	Kind of relationship	Place of the study	Source
Age	Younger people more mobile, more often leave rural areas	France	(Baccaïni & Dutreuilh, 2007; Détang-Dessendre et al., 2008).
Received education level	High-educated people migrate more often and more to cities	France	(Détang-Dessendre et al., 2008)
Type of employment	Self-employed works and executives migrate more than average individual, blue-collar workers migrate less than average individual	France	(Détang-Dessendre et al., 2008)
Fertility rate	Higher fertility in rural areas	Finland	(Kulu, 2013)
Family composition	Increased migrations when family composition changes (new couple, birth of child,…)	France	(Détang-Dessendre et al., 2008)
Social Network	Strong social network in place of resident prevents emigration	United Kingdom	(Zhang & Jager, 2011).

1.4. Abrupt and unexpected changes in population dynamics

An important consideration is that sometimes external factors, such as climate and climate change, hazards like earthquakes, flooding and volcanos and exceeding of the land's carrying capacity can trigger abrupt changes in population dynamics (Hunter, 2011). The same is true for social events like war or diseases. Noteworthy examples of events impacting the population dynamics in the past are listed in table 4.

__Table 4:__ Examples of events abruptly impacting population dynamics.

Description of event	Consequences related to population dynamics	When?	Source
Exceedance of the carrying capacity of the natural resources in Easter Island	Collapse of the population	Around 1500	(de la Croix & Dottori, 2008).
World war I and world war II in Europe	Combined more than 70 million civilians killed in the two wars, areas destroyed, millions of people displaced	1914-1918 & 1940-1945	(Kesternich et al., 2014; Mougel, 2010)
Tsunami in the Indian Ocean	>170k deaths Many areas uninhabitable ,>500k people displaced Increased fertility after tsunami	2004	(Nobles et al., 2015).
Earthquake in Haiti in 2010	65k-300k deaths About 42 million people displaced Population decrease of 23% in capital city Port-Au-Prince	2010	(Lu et al., 2012).
Civil war in Syria	Hundred thousands of deaths Millions of people displaced (national and international)	2011-present	(United Nations, 2016)

Those are just a few of the sudden events that have impacted population dynamics. The examples illustrate that natural, social and political events can abruptly change population dynamics. The focus of this research is about rural population dynamics in Europe for the last 50 years, but the above examples illustrate that unpredictable events can abruptly trigger unexpected changes in population dynamics. Events that can be seen most clearly on the population curve of Europe during the last 1000 years are the Black Death (around 1350 or 600 BP), the Little Ice Age (from about 1450 to 1750, or 500 BP to 200 BP), and the Second World War (1939-1945). As visible (figure 9), the Little Ice Age was responsible for the sharpest decline in population in Europe in the last 1000 years. The Little Ice Age caused social disruption, economic instability and food scarcity in Europe leading to migrations in the search for food and deaths because of food scarcity and social unrest (Pfister & Brázdil, 2006).

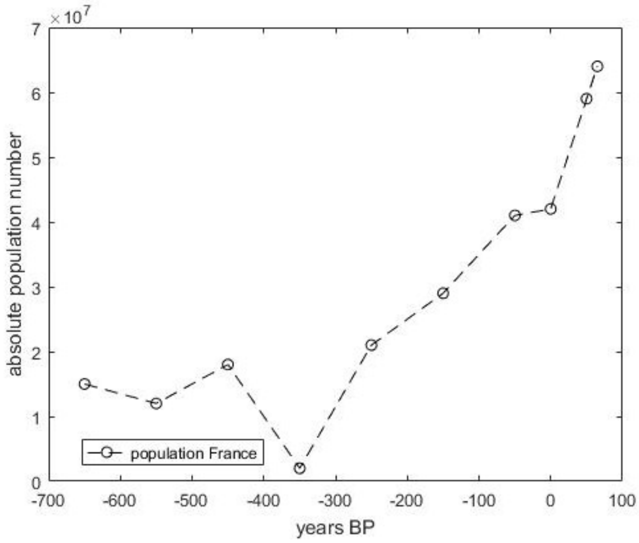


Figure 9: Evolution of the French population since 700 BP (own processing based on INSEE data).

1.5. Relevance of rural population dynamics

There are several motivations to invest in research on the mechanisms of population dynamics. Land Cover distribution, environmental issues and ecosystems and transportation networks are all affected by population dynamics and the spatial pattern of population settlements. Highways or secondary roads will not be built if there is not a large enough group of people that will use them and it makes also no sense to build a big shopping center in a totally remote area where nearly no people live. Reduced population, diminishing of rural villages and increased population and increasing-built-up area in other rural villages can disturb ecosystem functions and natural habitats, lead towards increased landscape fragmentation, increased flooding risks, and increased human trafficking. However, some consequence of rural depopulation can also be regarded as positive: in declining rural villages where the main cause of depopulation is farmland abandonment, the land can be reforested. The essence is that all those examples point on the importance of understanding population dynamics for making decisions related to spatial planning and understanding land cover dynamics.

Modelling population dynamics, one of the objectives of this research, can also provide a useful tool for taking policy decisions. The purpose of this research is not to judge which policy decisions should be taken, but to provide a tool for policymakers to evaluate the expected outcomes of possible policies regarding influencing population dynamics. The research can be useful for policy decisions at different scales. It can provide useful insights in which villages it can be strategic to stimulate employment, adapt transport networks, and enact policies to protect ecosystems. These processes are mainly related to regional or national governments, while this research can also provide useful information to the local governments of the villages, because it explores the effect of the services present in the villages on population dynamics.

Beside social relevance, understanding population dynamics is relevant from fundamental science point of view because it can lead to new insights in rural population dynamics and it can lead to new model approaches. Most studies of population dynamics examine the effect of one or a few factors and fail to develop a synthetic view taking into account both the characteristics of the population and the characteristics of the villages.

The advent of agent-based models (Parker & Meretsky, 2004) makes it theoretically possible to model population dynamics starting from the decisions of individuals. Despite this potential only a few studies explored these possibilities (Zhang & Jager, 2011) but all of them lack an external validation of the model results.

This work will address the research gap by evaluation the potential of agent-based models for the simulation of population dynamics in rural areas, whereby the Marne department in northern France is taken as an example application.

1.6. Research Questions and Research Objectives

Research Question 1: What is the spatial pattern of rural population dynamics ? What is the spatial pattern of villages with increasing or decreasing population? For example, are rural villages with declining population evenly spread over regions? Or are they clustered together? Can they occur everywhere or only far from regional cities or far from railway transport infrastructure? Are there noticeable similarities when compared to other characteristics of the area like land cover?

Research Question 2: What are the push and pull factors[2] (controlling factors) of rural population dynamics related to

a) **The typology of the population:** The purpose is to investigate how the characteristics of the population are linked to population evolutions. For example, wat is the influence of the age and the received education on the occurrence and typology of migrations.
b) **The typology of the rural villages:** Which characteristics of the villages can explain the evolution of the population in the village? Which characteristics are push factors of migration and which characteristics are pull factors of migration? Many variables will be explored, related to the EP of the villages, the number of services in the village, the remoteness of the village, the soil fertility and more.

[2] Push factors are the factors causing the people to leave their village of residence while pull factors are the factors that attract people to other, new places of residence.

Gathering those type of data and answering those research questions can lead towards new insights in rural population dynamics.

Research Question 3: What are the implications of those rural population dynamics for the characteristics of the villages and on how will the problem of rural abandonment evolve?

a) **What are the results of depopulation (or increased population) on the service level of the rural villages?** If the population falls or rises in rural villages, does this affect the number of services present in the villages (shops, schools, …) and what is the time lag between the population evolution and changes in the number of services in the rural villages?
b) **How will rural abandonment evolve ?** What is expected from those analyses how the problem of rural abandonment will evolve in the future? Can policy interventions influence the process of rural abandonment?
c) **Will there be other implications?** Are there land use changes expected because of the rural population dynamics, will the spatial planning change, …?

Research Question 4: Is it possible to simulate the key processes of rural population dynamics with an agent-based simulation model?

- Is it possible to simulate past dynamics?
- Is it possible to simulate future evolutions?
- How accurate are those models? (validation)

So, the main **research objectives** can be summarized as follow: Analyze the spatial pattern of the demographics of rural villages, relate the population typology and village typology to the population evolutions, and develop agent-based models and test whether those models are able to accurate simulate past population dynamics and simulate future population evolutions. Those research objectives will be investigated with a case study of the French Marne department. The next section provides arguments for selecting the Marne department as the study object.

CHAPTER 2: STUDY AREA

2.1. Study Area Justification

The Marne department belonging to the French region Champagne-Ardennes was selected as study area. Firstly because it is characterized by large rural areas of which many parts are relatively remote from urban centers. Secondly, the geographic and socio-economic settings of the remote villages in the Marne department are representative for many regions in the French countryside and can also be found in other European countries such as Spain, Romania, Italy and Slovenia. These settings are quite different from the settings in the e.g. Belgian countryside in which all villages, even in the most remote places are relatively well connected with smaller or larger cities. According to the European landscape classification is nearly all of Flanders classified as suburban area (Brandmülle et al., 2016).

Thirdly, the Marne department is characterized by an important variability in rural population dynamics: some rural villages have been losing population for decades while some have an increase in population and others show a more complex evolution with periods of increase and decrease.

And fourthly, the reason for selecting France is that there are practical limitations because it is crucial to conduct some weeks of field work for this research, and there's also the matter of understanding the language for conducting. The language during this field work was French which the researcher is capable of understanding. Especially in rural areas, the knowledge of English in France and other European countries is very low.

2.2. General Description of the Study Area

2.2.1. Location and context of the Study Area

Figure 10 shows the location of the Marne department. France is administratively subdivided into regions (n=13) and departments (n=96). The Marne department is part of the larger region of the Champagne-Ardenne. Within the departments there are arrondissements, cantons and villages (government France, 2017). The villages are the lowest administrate level on which this study will focus.

The Marne department has an area of about 8162 km² and contains 617 villages. About 570 000 people live in this area (2013), corresponding to a mean population density of 70 people per squared kilometer. (government France, 2017; INSEE, 2017). The most important cities in the department are Châlons-en-Champagne, Epernay and Reims.

Between 1968 and 2012, most of French departments gained population with a strong population growth of over +30% in more than 30 of the 96 departments. 11 of the 96 departments, all situated inland, have lost population between 1968 and 2012 (figure 10). Departments along the Atlantic and Mediterranean coasts show an increase in population, which suggests the possible impact of physical landscape characteristics on population evolutions.

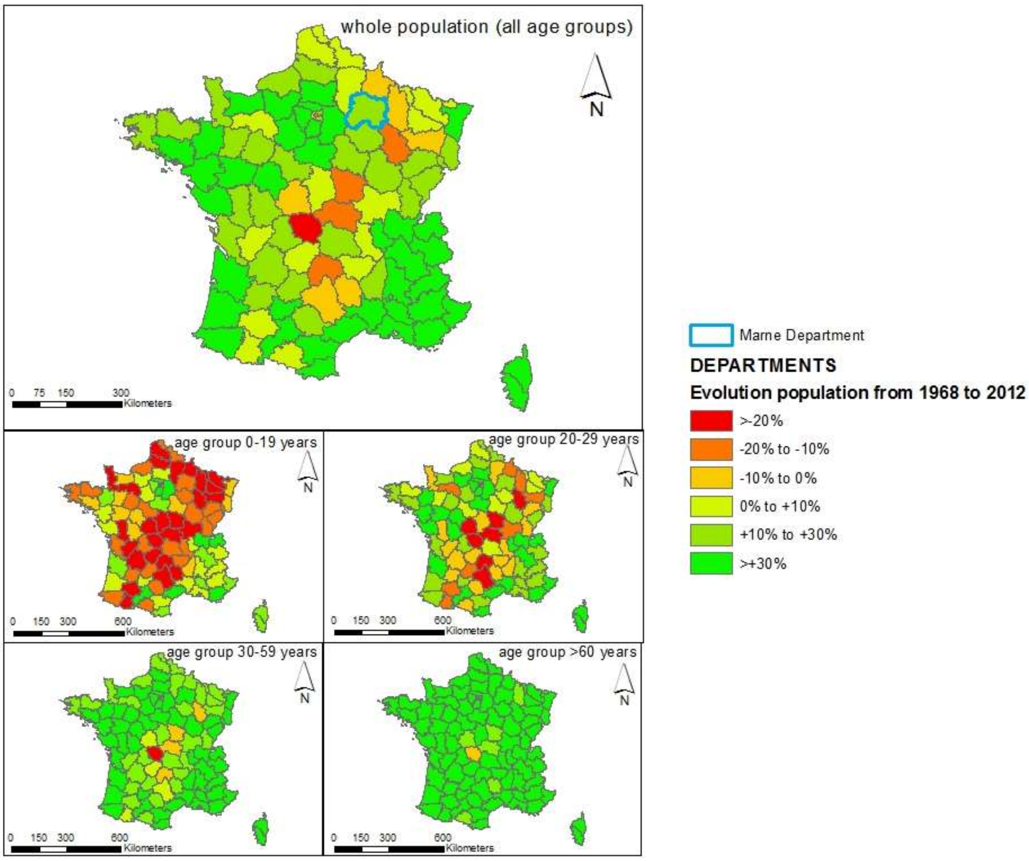


Figure 10: French departments with location of the Marne department indicated. Relative population evolution from 1968 to 2012 is shown for the entire population and for different age groups. Maps are developed (own processing) based on row data from (INSEE, 2015).

The Marne department had a slight to moderate population growth between 1968 and 2013. For this study, we will consider the period 1982 to 2013[3]. In France, the population growth between 1982 and 2013 was +24.1% while the population growth in the Marne department was +4.5% for the same period (INSEE, 2017). The evolution of the number of people living in a department results from natural population growth (number of births tops number of deaths) and from migrations (equation 1) . Migrations can be intra-departmental, inter-departmental or international. The number of international migrations are low compared to the inter-departmental and intra-departmental migrations inside of France. There is a net positive migration towards France with Portugal, Italy and Spain being the most important sources of migration (INSEE,2015).

[3] Because of availability of data

It is important to understand that natural growth also contributes significantly to population fluctuations. Because the number of births greatly outweighs the number of deaths in France and the Marne department (figure 12), the birth/death ratio (figure 11) is >1, indicating a net natural growth of the population. The last 10 years, the birth/death ratio for the Marne department has been equal to the French average, but before 2005, the birth/death ratio was clearly higher for the Marne department than for France as a whole. Despite this, the increase in population in the same period was seriously lower in the Marne department compared to the French average, which implies the natural growth of the population in the Marne department is to a large extent compensated by net emigrations from the Marne department to outside the department. This is confirmed by demographic reports (Baccaïni & Dutreuilh, 2007; INSEE, 2010). The net inter-departmental migrations (migrations from outside France not considered) from and to the Marne department were under -0,34% for the period 2001-2006, indicating more people leaving the Marne department than moving to there (INSEE, 2010). The average inter-departmental migration rate within France is about 2.5 to 3%, while the average inter-communal annual migration varies between 6 and 8%, implying that the majority of the migrations in France are intra-departmental (Baccaïni & Dutreuilh, 2007). The residential mobility is far higher for younger people between the ages of 20 and 29 years compared to the other age groups and the residential mobility decreases further with increasing age (Baccaïni & Dutreuilh, 2007).

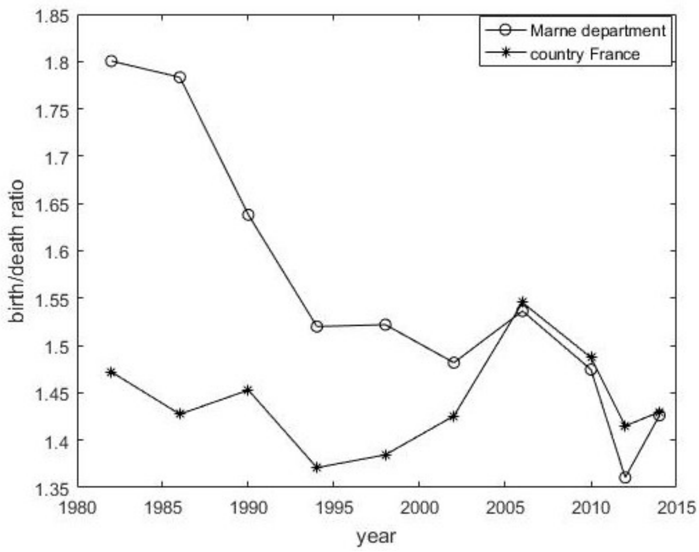
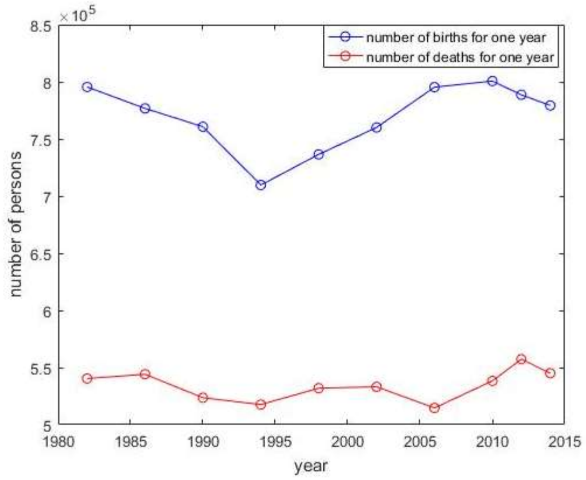


Figure 11: Evolution of birth/death ratio since 1982 for France as a whole and for the Marne department. Own processing based on row data from (INSEE, 2015).



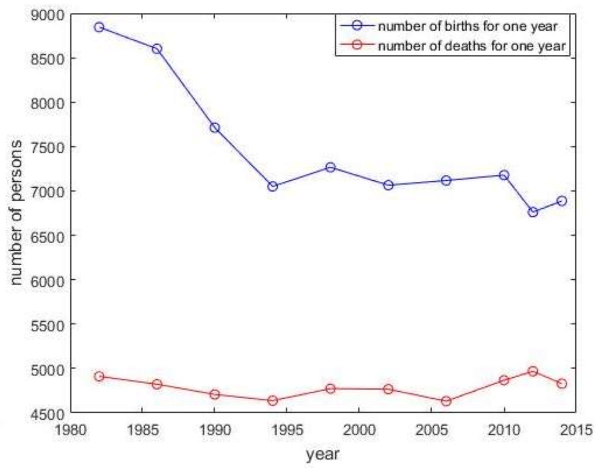


Figure 12: Evolution of number of births and deaths for France (left figure) and the Marne department (right figure). Own processing based on row data from (INSEE, 2015).

2.2.2. General description of the Study Area

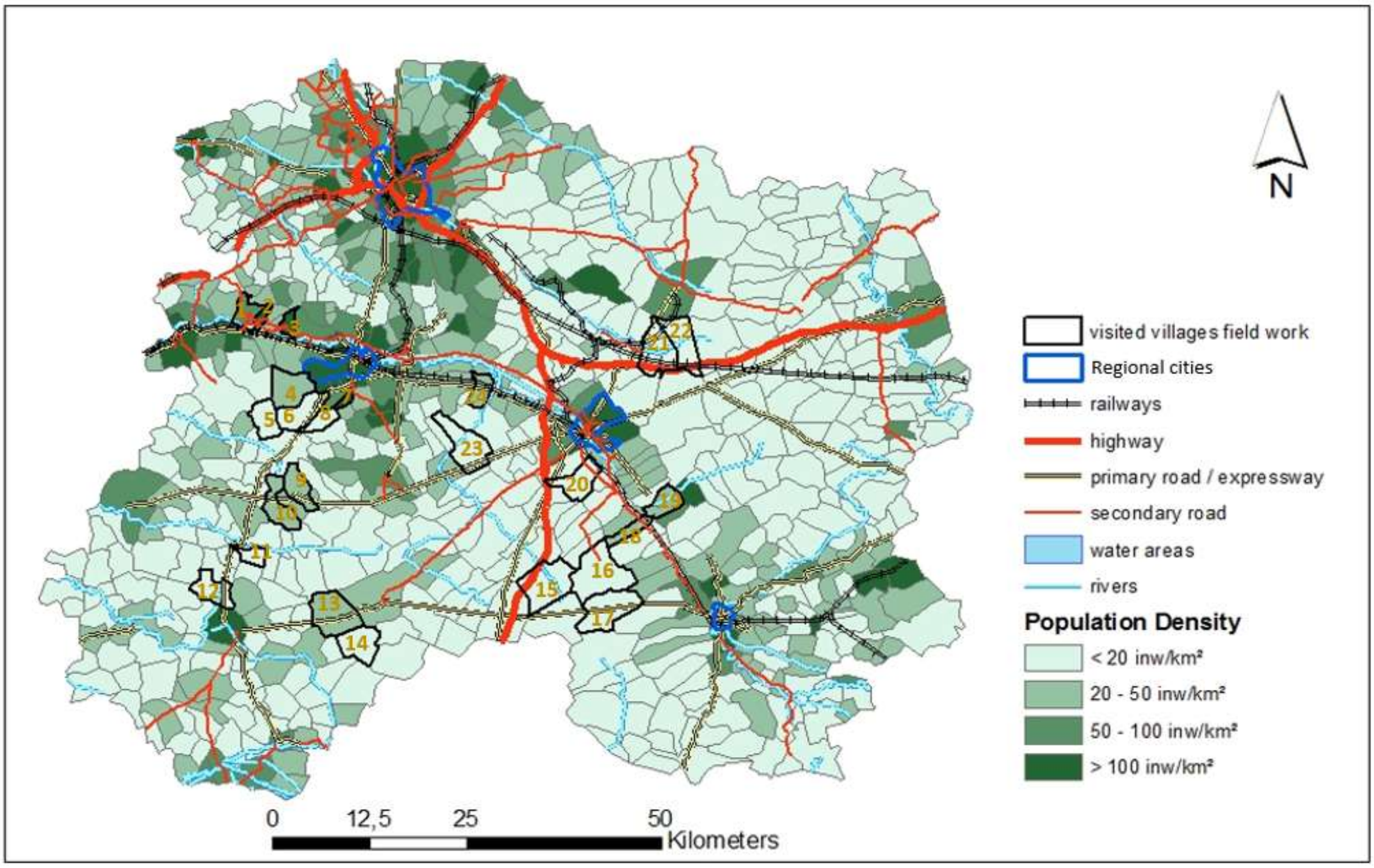


Figure 13: Characteristics of the study area like the population density, transport networks and location of regional cities and villages visited during field work. Own processing using data from INSEE and shapefiles from ESRI. The visited villages during field work are labelled from 1 to 24. 1=Châtillon-sur-Marne. 2= Binson-et-orquigny. 3=Venteuil. 4=Saint-Martin-d'Ablois. 5=Le Baizil. 6=Brugny-Vaudancourt. 7=Monthelon. 8=Morangis. 9=Etoges. 10=Congy. 11=Oyes. 12=Lachy. 13=Connantre. 14=Corroy. 15=Dommartin-Lettrée. 16=Faux-vésigneul. 17=Coole. 18=Vitry-la-Ville. 19=Pogny.20=Ecury-sur-Coole. 21=La Cheppe. 22=Bussy-le-chateau. 23=Pocancy. 24=Jâlons.

In most of the Marne department the population density is relatively low with less than 20 inhabitants per squared kilometer. Denser populated areas can be found in the cities and their surroundings. The major roads and the railways mainly connect different cities (figure 13).

2.3. Physical characteristics of the Study Area

2.3.1. Climate

The region is characterized by climate type “Cfb” according to the Köppen-Geiger climate classification, implying an temperate oceanic climate without dry season and with a warm summer, with the coldest month averaging above 0°c, an average temperature below 22°c for all months, at least four months where the average temperature is above 10°c and no significant difference in precipitation between the different seasons (Kottek et al., 2006).

2.3.2. Geology

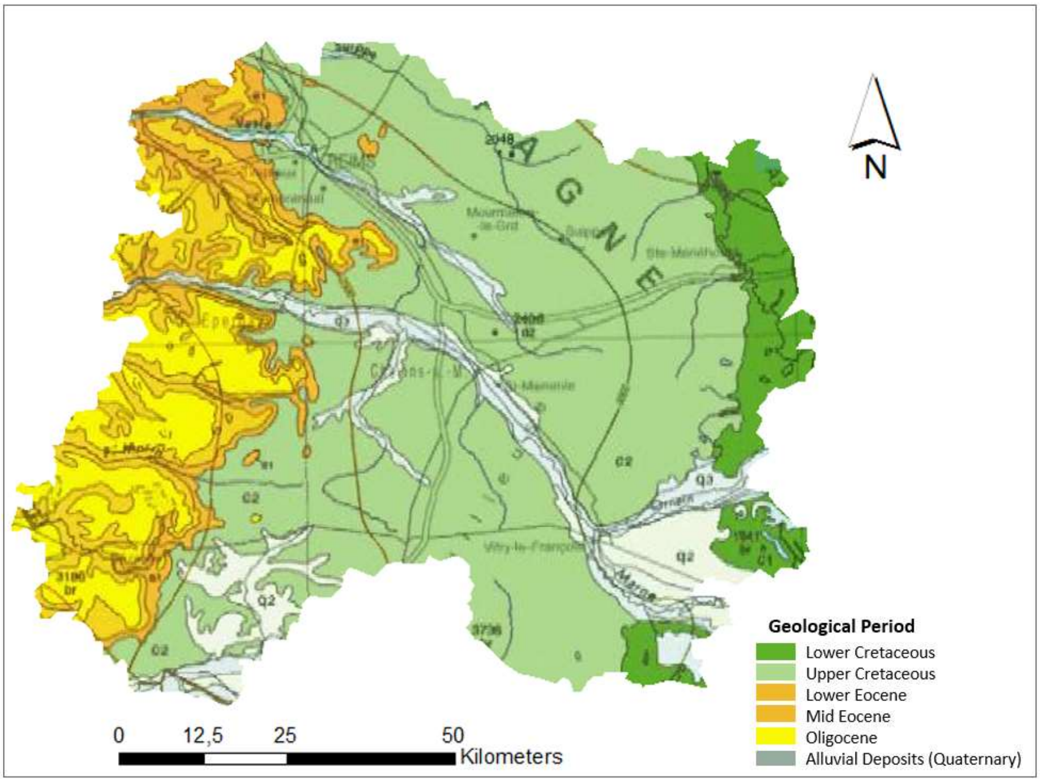


Figure 14: Geological map of the study area, adapted from (BRGM, 2017).

A comparison of the geological map with the other maps shows clearly that the most big cities and the majority of the important transportation axes (roads and railways) are located on the recent alluvial Quaternary deposits (figure 13, figure 14). This area with quaternary deposits corresponds to the area with lowest elevation, along the majors rivers (figure 15).

__Table 5:__ Overview of the geological formations from various geological times in the Marne department

Geological Period	Time period
Lower Cretaceous	145-100Ma
Upper Cretaceous	100-66Ma
Lower Eocene	56-41Ma
Mid Eocene	41-38Ma
Oligocene	34-23Ma
Quaternary Alluvial Deposits	<2.5 Ma

2.3.3. Elevation

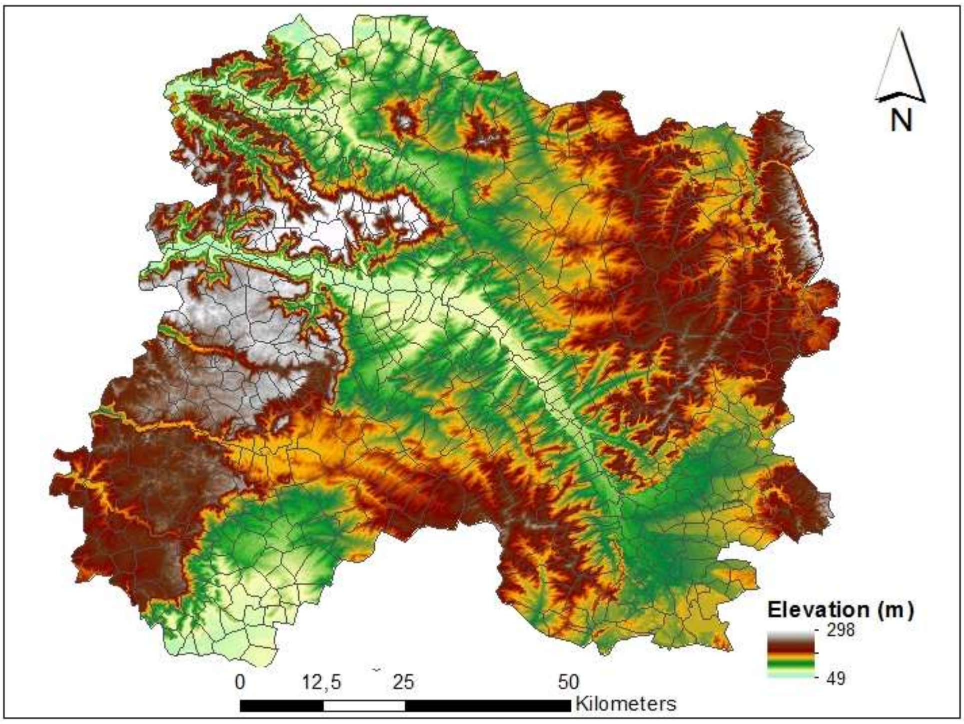


Figure 15: DEM of the study area, own processing using SRTM elevation data (NASA, 2017).

The DEM is developed from elevation data from the 2000 Shuttle Radar Topography Mission (SRTM), developed by NASA. The DEM has a resolution of 90m (NASA, 2017). The DEM shows the spatial pattern of the elevation. The elevation in the Marne department ranges between 49m and 298m above sea level. The long, small area with the lowest elevation is the valley of the Marne River. A visual comparison with the land cover map (figure 16, figure 17) shows that forest areas are always located at the higher elevations in the landscape. Arable land can occur on the plateaus or at lower elevations. In general, low elevation areas are more densely populated than high elevation areas (figure 13, figure 15).

2.3.4. Land Use

The land use maps in figure 16 and 17 are derived from the Corine Land Cover maps. The 1990 map is based on Landsat-5 MSS/TM satellite data while the 2012 map is based on IRS LISS III and Rapid Eye dual date satellite data. In both cases, the spatial resolution is resampled to 100m (Manakos & Braun, 2014). Beside satellite data, ancillary data like aerial photographs and topographic maps were used (Eea, 1994). Both maps have a thematic accuracy of at least 85% (Manakos & Braun, 2014). The original Corine Land Cover map contained an enormous amount of land cover classes in the Marne department (44 classes), therefore the maps were reclassified using the seven most important land use classes. Although the Corine map is called a land cover map, it is in fact more like a land use map, because land cover refers to the physical land type, like water or impervious area, while the land use refers to how people use the land, like arable land and industry (NOAA, 2017). However, there is often a lot of confusion between the two terms and often those terms are used as interchangeable terms (Comber, 2008).

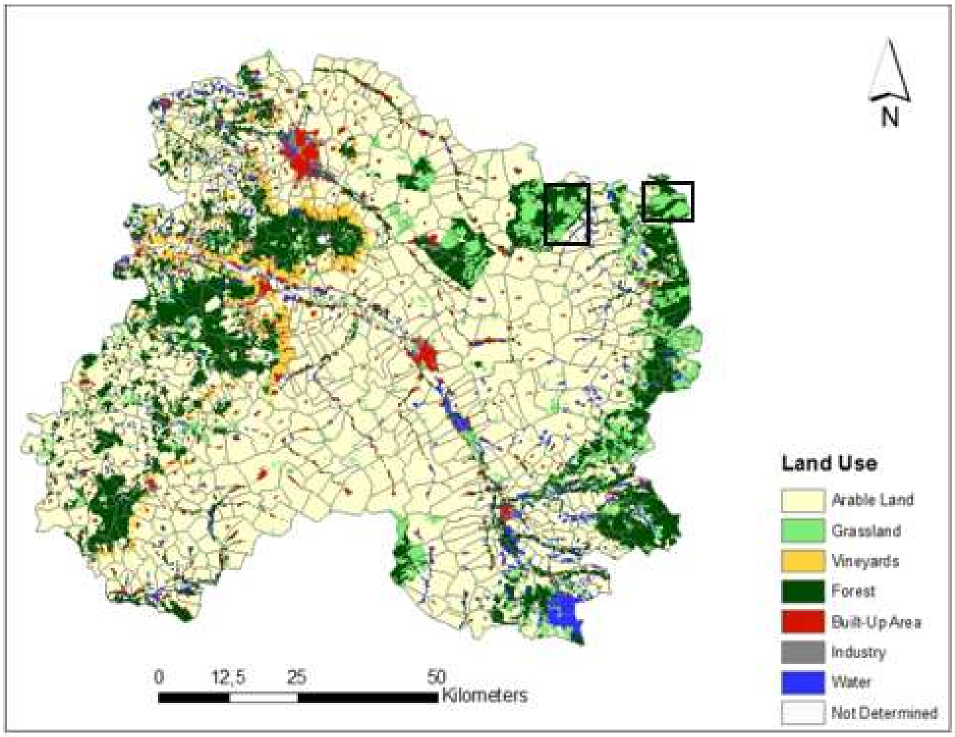


Figure 16: Land Use Map of the study area, 1990 (adapted from Corine Land Cover 1990).

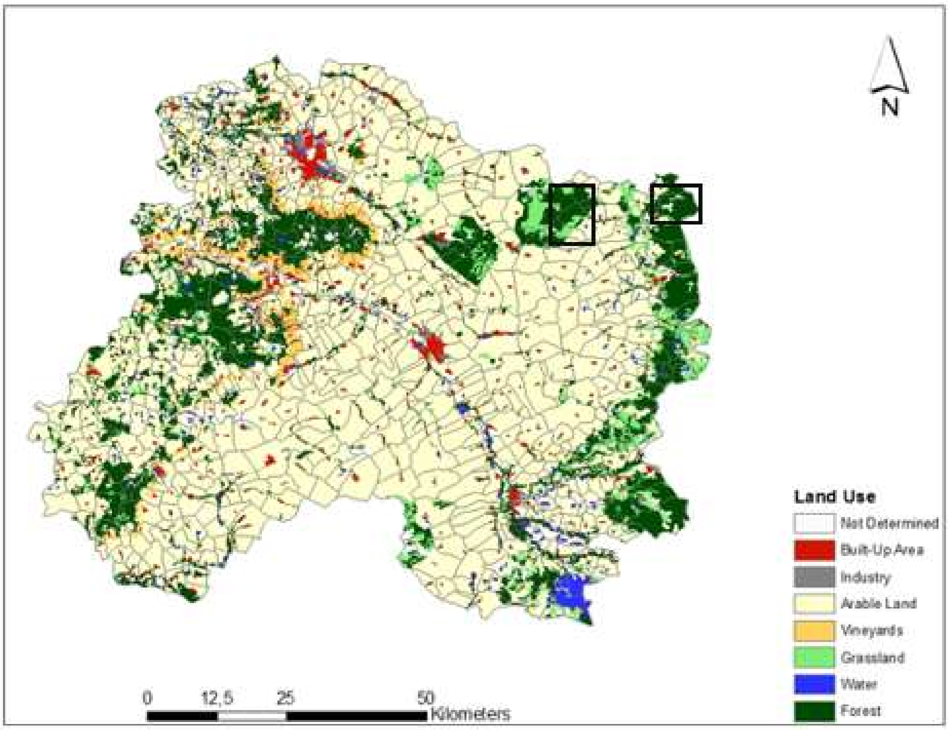


Figure 17: Land Use Map of the study area, 2012, adapted from Corine Land Cover 2012. The two black boxes indicated on the figure are areas where there is a visible change in land use from 1990 to 2012. In those areas, a significant amount of grassland is converted to forest.

On the land use maps of 1990 and 2012, arable land is the dominant land use in the Marne department (figure 16, figure 17). Most of the forested areas are characterized by low population densities (figure 13, figure 16, figure 17). Rural villages with arable land as land use can both have a low or a high population density (figure 13, figure 16, figure 17) and logically, the largest patches of built-up area correspond to the cities, which have the highest population densities.

A pixel-by-pixel analysis of the change in land use between 1990 and 2012 shows a decrease in grassland and pasture and a small decrease in arable land. The forest area increased during this period (figure 18). The greatest shift in land use is the conversion of large parcels grassland and pasture to forest. Those changes are visible on the land use maps and are indicated within the black boxes on the figures (figure 16, figure 17). The forest has area increased from 17.3% to 17.9% from 1990 to 2012 while the area pasture and grassland has decreased from 7.0% to 5.9% in the same period (figure 18).

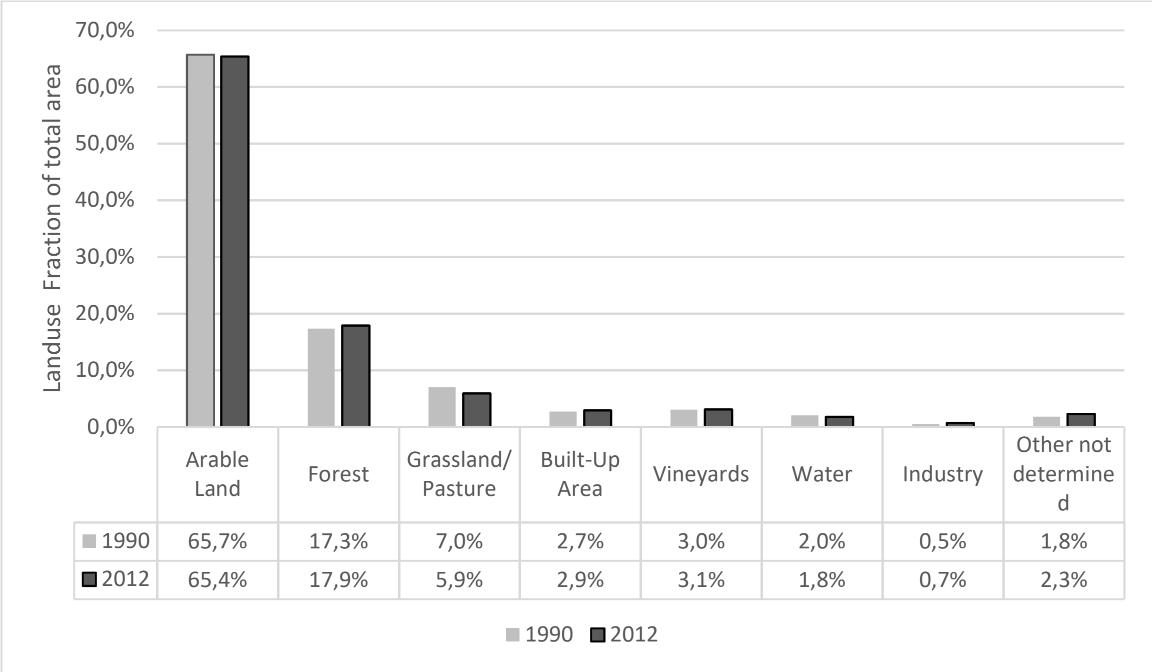


Figure 18: Land Use Distribution in the Marne department in 1990 and 2012. Own processing using data derived from the original Corine Land Cover maps.

2.4. Demographic analysis of the Study Area

Figure 19 shows the population evolution from 1968 to 2013 for the individual villages of the Marne department. The Marne department is characterized by a high rural population compared to the European average. But this is highly variable within the department: while some regions such as the region around Reims have an urbanization level higher than 80%, in some other regions like the southwestern and the eastern part of the Marne department, the urban population is less than 50%. On the level of individual villages, there are villages for which the population has been increasing or decreasing continuously over the considered time period, but there are also villages with alternating periods of population increase and decrease (figure 20).

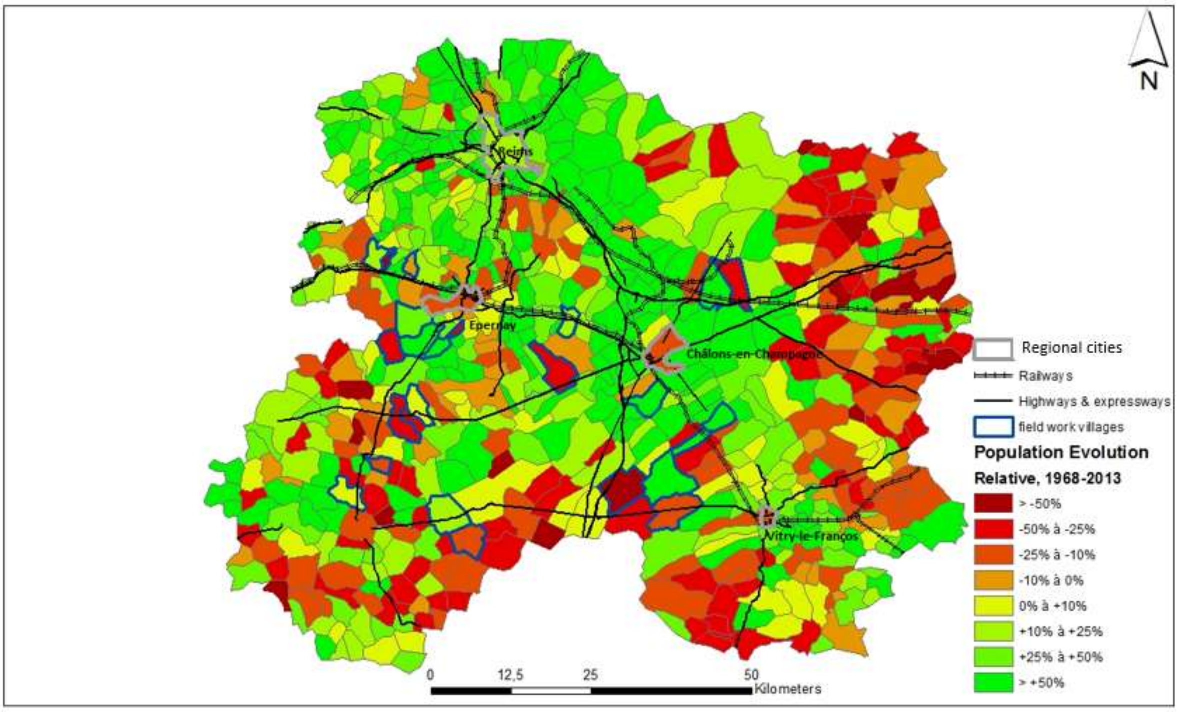


Figure 19: Evolution of the population in the Marne department, 1968-2013. Own processing based on INSEE population data

The share of people living in urban areas in the Marne department was decreasing circa linearly from 1968 to 2013 while the population number in suburban areas was increasing continuously in the same period. The percentage of people living in rural villages remained almost constant during the considered period (figure 21). Urban population has decreased in three of the four largest cities in the Marne department: Châlons-en-Champagne, Epernay and Vitry-le-François. Only the biggest city of the department, Reims, has an increasing population (figure 20 and figure 22). Most suburban areas (indicated on figure 20a)[4], have an increasing

[4] How is determined which villages are rural, urban and suburban is comprehensively discussed in the methodology chapter.

population, especially in the suburbs around Reims. Only the suburban villages around Epernay show a decrease in population (figure 20). For the rural villages, the population trend is highly variable: several rural villages have an increasing population while several other rural villages have a decreasing population.

From 1968 to 2013, the population increased in the northwestern and the central part of the department while predominantly the eastern and southern part of the department were losing population (figure 19). It is remarkable that the regions with decreasing population have low population density (figure 13, figure 19). The main causes for rural depopulation in the southern and eastern part are a decrease in population aged under 30, however, some villages in the southern and eastern part were losing population in all age groups, including the oldest age groups of >60 year for which there is an increase in population in the most other villages (figure 22). The decrease in people aged under 19 years can be partially explained by the decreased number of births until 1994 in the Marne department (figure 12). After 1994, the number of births remained stable (figure 12). Another fact is that the birth/death ratio is higher in the rural villages than in the urban and suburban areas due to higher fertility in the rural villages. For example, the birth/death ratio was on average 1.47 in 2013 in the Marne department and the average for the urban areas was 1.42 while the average for the rural areas was 1.53 (calculated based on INSEE data). Higher fertility can be expected in the rural villages because people are on average poorer and less-well educated (Azarnert, 2014). This means the natural population growth is higher in the rural areas than in the urban and suburban areas.

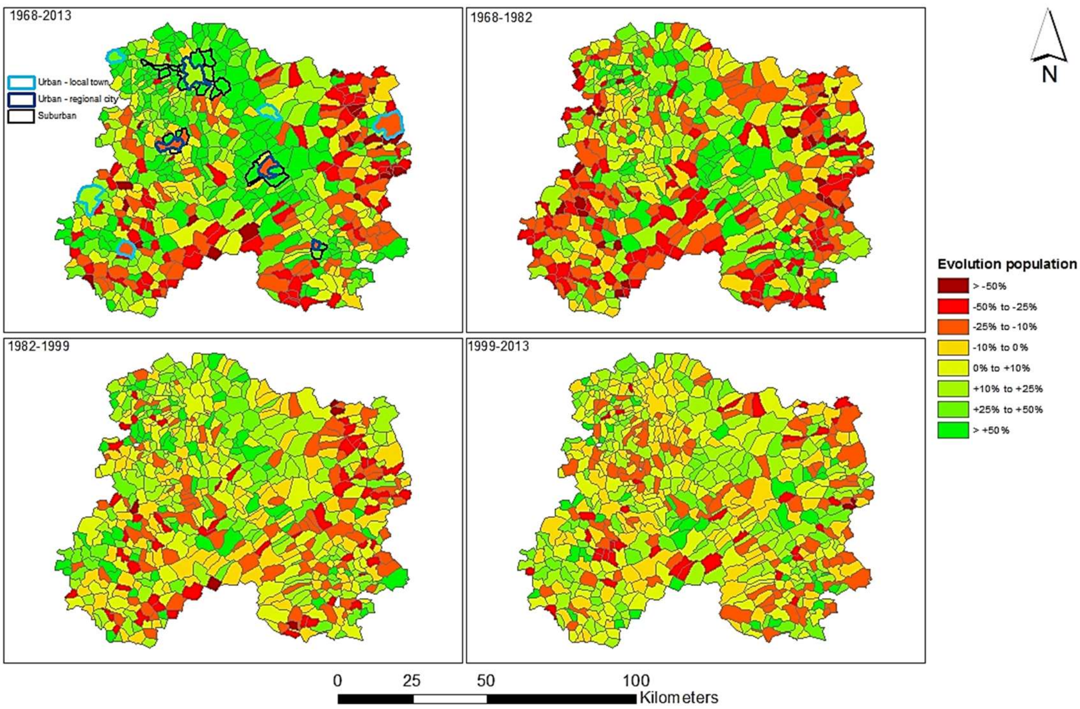


Figure 20: Evolution of the population in the Marne department from 1968 to 2013 and for some subperiods. Own processing based on INSEE population data.

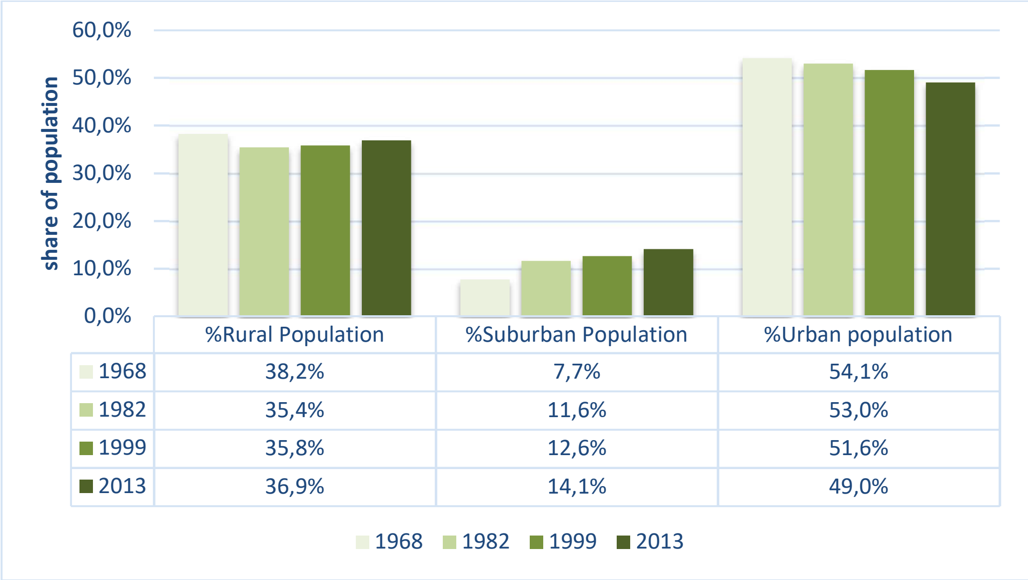


Figure 21: Share of rural, urban and suburban population in the Marne department. Evolution from 1968 until 2013. Own processing.

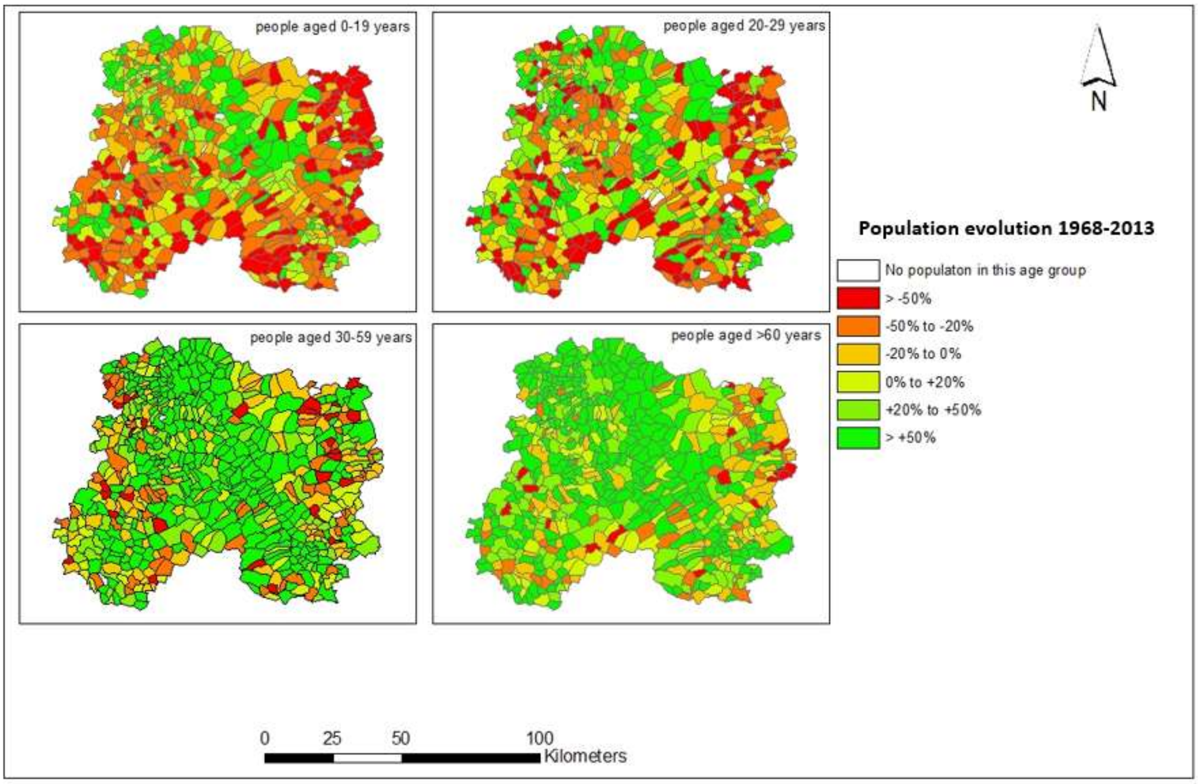


Figure 22: Evolution of the population from 1968 to 2013 for people of different ages. Own processing based on INSEE population data.

CHAPTER 3: SPATIAL AND STATISTICAL ANALYSIS OF RURAL POPULATION DYNAMICS

3.1. Data and methodology

3.1.1. Defining rural and urban

The focus of this research is as mentioned on rural population dynamics with a focus on the Marne department. Therefore, it is important to define which communes can be identified as rural villages and which communes are urban or suburban zones. The identification of cities and suburban villages is not always straightforward. Various methods exist in literature to identify those zones (Forsyth, 2012). Most definitions are based on location, population density and newness, but there are also a lot of other less frequent used definitions such as classification on activities, sociocultural characteristics or even the PM 2.5 level, which is an indicator for air pollution based on the amount of particular matter (Forsyth, 2012; Zhou et al., 2004). Delineation of urban and suburban zones is in that sense a controversial topic, but for this study delineation is needed because the rural population dynamics are the focus in this research.

In this study, four different settlement types will be delineated: regional cities, local towns, suburban villages and rural villages. What follows is a short discussion of the criteria used to identify each of the four zones.

<u>Urban – Regional city :</u>

The French government definition of a regional city is used in this research. According to the French government, a regional city has at least 1500 employment places in the city center, has several municipalities without pockets of clear land surrounding the city center and at least 40% of the employed resident population works in the city center or the suburbs (INSEE, 2016).

<u>Urban – local town:</u>

According to the definition of the French government, a village can be considered urban if more than 2000 people live there and there is no space gap larger than 200 meters between buildings (INSEE, 2016). The latter criterion is impossible to check with the available tools and data for this research. Therefore, the definition used in this research is a combination of the first criterion used by the French government (2000 inhabitants) complemented with the definition used by the OECD stating there should be a population density of at least 150 inhabitants per squared kilometer (INSEE, 2016; OECD, 2016). So the criteria used in this research to identify a local town are:

- At least 2000 inhabitants
- A population density of at least 150 inhabitants per squared kilometer
- Does not belong to the class regional city.

<u>Suburban village:</u>

In this research the delineation will be carried out based on following criteria:

- A suburban village should be located within a reasonable commuting distance from the urban center, with a threshold of 25 km chosen here.
- A suburban village should have a population density of at least 150 inhabitants per squared kilometer.

Many studies and reviews about suburban areas suggest the commuting distance and population density as criteria, however, there is no consensus of the specific threshold of those indicators (Forsyth, 2012).

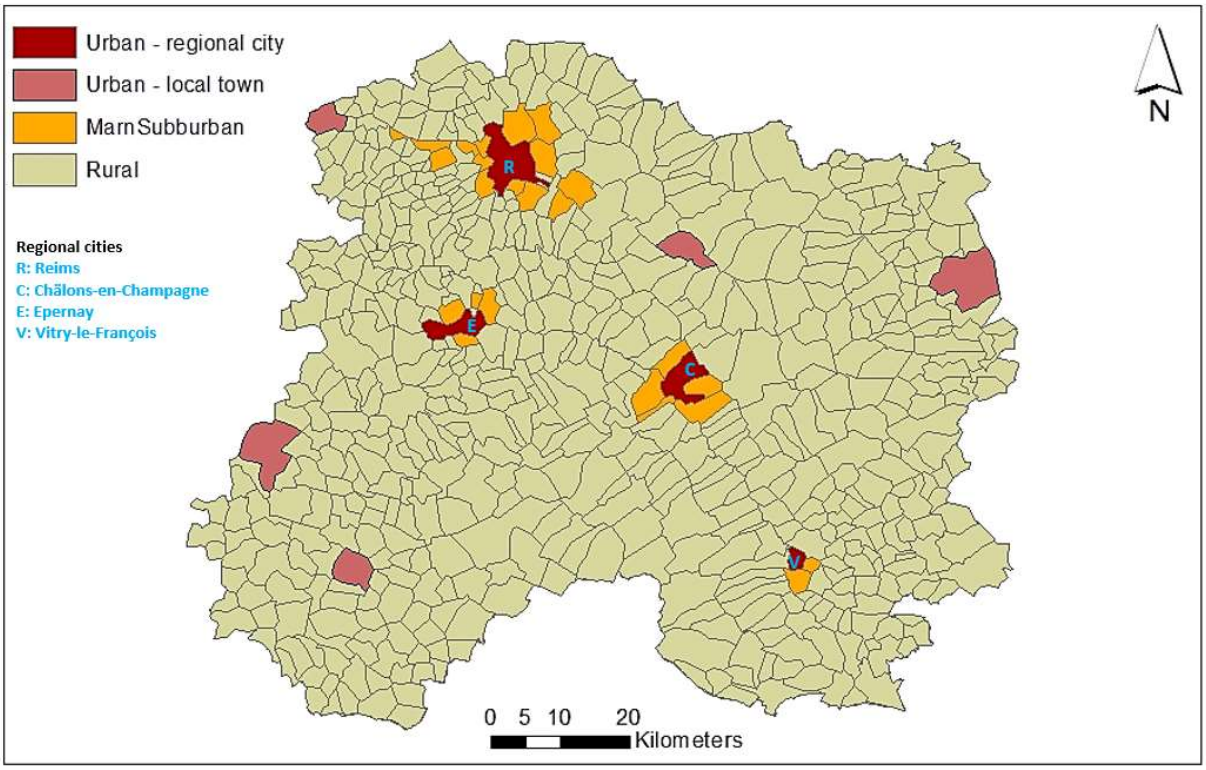


Figure 23: Classification of urban, suburban and rural villages in the Marne department. Own processing.

Besides these criteria, a visual inspection of the villages surrounding the regional cities was conducted. If a village, classified as rural under the proposed definitions, seems to be more suburban in reality based on a visual inspection in Google Earth, the classification was manually overruled. An example of this is Saint-Martin-sur-le-Pré in the suburbs of Châlons-en-Champagne. The village is clearly an extension of the city (figure 24). But because the built-up area of this village is only a small fraction of the village, the village also includes large wheat fields without residential zones, and the village has a low population density, it was not initially classified as suburban. Therefore, a manual intervention corrects this situation and classifies the village as suburban. The application of manual interventions was very limited (n=3), in the majority of the cases the classification was correct following the proposed definitions.

Figure 24: Example of a village wrongly classified as rural following blind application of classification rules . In this case, a manual correction is executed. The borders of the considered village, Saint-Martin-sur-le-Pré are shown on the figure. Adapted from (Google Earth, 2017).

<u>Rural village:</u>

All villages that are neither urban or suburban.

The result of the followed classification in urban (regional city and local town), suburban and rural is presented in figure 23. An aspect important to mention is that some villages that where not yet suburban in 1968 can become suburban over time. What is visualized on figure 23 are all villages that are suburban in 2013, even if the villages were not yet suburban in 1968.

3.1.2. Field Work

3.1.2.1. Preparation of terrain analysis

Several actions preceded the field work. First, information was obtained related to rural population dynamics and the problem of rural depopulation. This information was obtained by analyzing existing literature regarding population dynamics related to controlling factors of population dynamics and literature regarding cases of demographic evolutions across the globe. An introductory demographic analysis and introductory correlation analyses are carried out. Those analyses were used to decide which questions are appropriate to be asked during field work. The results of the exploratory correlation analysis are not shown. Later on, a more comprehensive correlation analysis, including field work data, is shown.

Existing literature and the exploratory analyses provide a rich source of information, but many aspects regarding population dynamics are under-investigated and there is often insufficient data regarding migration motivations of individual households to accomplish in-depth analyses related to the subject. Therefore, it was considered as very useful to collect a lot of data on terrain which are not available or accessible through other sources such as existing databases. A separate section (3.1.3) describes the data obtained from field work.

Moreover, through conducting terrain work, it was possible to extract crucial information about how, in which manner and in which context, data obtainable from other sources publicly available could be used. An example of this is housing prices (HP). Those are available from existing databases, but through questioning the local people, it became clear how the relation between this variable and the evolution of the population should be interpreted. Globally, there is a net slightly positive correlation between HP and the evolution of the number of people in a village, but those data mask the complexity of this variable. Part of the population, mainly population with a low income, will migrate to places were the HP are low , and often for those people, the HP are a more serious concern than the services present in the villages. On the other hand, for people with a sufficient income, the HP are often from less importance compared to the EP and the SL of a village. This kind of complex relationships between population evolution and variables becomes noticeable by questioning people in the field, and are not visible in the raw data available from existing databases.

Finally, the field work is also important to gain better insights in the process of rural population dynamics and rural depopulation. By spending some weeks in the rural villages and by interacting with the local population, insights are generated that cannot be generated through statistical data analysis.

Before conducting the field work, a comprehensive interview question list was drafted. The interview language was French. First, there are several questions related to characteristics of the interviewed people and their relatives (children, spouse, parents, sisters, brothers). Subjects of the questions were: their age, highest received level of education, type of employment, location of employment and their migration path. The migration path can be defined as a summary of all places where a person has lived since his/her birth. Second, questions were asked about the number of services present in the village and the evolution of the number of services. For example, the question was asked how many mini-markets there are present in the village, and if there were minimarkets that started up or went out of business during the last 50 years. Thirdly, there is a list of questions regarding which factors the questioned people find important for living at a certain place. The question list can be found in Appendix A.

For the development of the interview question wording, a review paper about preparing for interviews and developing interview protocol frameworks was used (Castillo-Montoya, 2016). In the first step, after developing a preliminary questionnaire, a check was carried out whether the questions are in line with the data that are wanted to obtain in order to facilitate answering the posed research questions of this study. To check the usability of the interview protocol, some test interviews were performed and some people were asked to give feedback or their opinion about the interview protocol. Receiving feedback on interview protocols and piloting the interview protocol on a test sample are important steps in developing a final interview protocol (Castillo-Montoya, 2016). About 15 people living in rural areas in Belgium took part in the questionnaire. The population dynamics of rural areas in Flanders cannot be easily compared with the rural villages in French, however, conducting test-interviews provides a useful tool to detect inconsistencies and missing elements and they also enable you to find out whether the questions are clearly formulated and whether the respondents understand what is meant with the question. It also indicates whether the order of the questions is logic (Castillo-Montoya, 2016). This way, the test interviews make it possible to iteratively improve the interview list. The pilot interviews were conducted in June 2016.

The interview questions are mostly quantitative based to obtain data. Most of the questions asked are factual and not open to interpretation. If a person claims to be 50 years old, has a university degree and is employed 30 km from his residential place as engineer, has two children of respectively 27 and 30 year who are migrated to suburban areas a few 100 km from the place the interviewed person lives, and claims to live in a village where there is one butchery and zero bakeries in the village because the last one closed two years ago, this are facts. Of course, it is impossible to check whether the interviewed person is telling the whole truth, but the results are believed to be independent of the researches. One more subjective part of the interview is to ask people how important they find different services in the villages. These were closed questions implying the answers are not open for interpretation. If another researcher asks the same questions, it is plausible to assume that he/she will obtain the same answers as another researcher.

The selection of the study area (Marne department) is explained and discussed in chapter 2. The villages were selected with the aim of collecting a representative sample of the rural settlements in the Marne department. Population evolution was the main criterion for the selection: the sample of 24 villages include villages with continuously declining and continuously increasing population and villages with a more complex population trend (with phases of rise and decline during the studied time span). A second criterion was the landscape type: both villages on the Brie-plateau and in the lower lying chalk-plain (Dry Champaign) were selected. The selected villages are shown in figure 13. In total, 24 villages were selected to visit during field work. A practical limitation is that all the field work is conducted by bicycle transportation, limiting the amount of distance possible to travel. There was no operational public transport in the regions visited. For that reason, often adjacent villages are visited.

3.2.2. Conducting the terrain analysis

The field work was conducted from July, 12, 2016, until August, 1, 2016. In these 19 days, about 150 hours of field work are conducted, not taking into account transportation times. In total 24 villages were visited. The different villages were visited by bike, which resulted in a total bike distance of about 800km, corresponding to a travel time of about 40 hours.

During field work, each day one to three villages were visited. In each village, the following activities were carried out:

- Semi-structured interviews with local households. In total, 171 interviews were conducted.
- For some villages, the mayor or an alderman of the village was interviewed. In six villages, there was a conservation with either the mayor or one of the aldermen.
- For each village an inventory of the number of the services present (schools, shops, bakeries, …) was made.
- For each village a small report was drafted describing the landscape, the condition of the houses and the state of the villages in general.

The interviewed households in the villages were selected following a stratified-random sampling procedure, whereby the different quarters (or streets) of the villages were visited on foot and within the quarters the households were (semi-)randomly selected by ringing at the door. For quantitative analysis, random sampling is rigorous and often considered the best option (Marshall, 1996).

The response rate was in general rather low. In total about 600 people were asked to participate in interviews but circa 75% refused. This resulted in a total number of 171 interviews during which information on the time-space path of 400 persons was collected, since people were also asked about the migrations of their children. A reason for the low response rate could be that the interviews took place in the days following a terrorist attack in Nice (July 14, 2016). The response rate varied between the villages, with in some villages the response rate reaching 70% while in other villages barely 10% of the people agreed to participate. In general, elderly people more often refused to participate than younger people.

An analysis of the population pyramid of the rural villages of the Marne department (figure 25 left) and the population pyramid of the interviewed population (figure 25, right) shows that the interview sample seems quite representative for the rural population of the Marne. The population pyramids shown on figure 25 are based on the rural population excluding the population in the suburban and urban zones[5].

A possible source of bias in the data collected during the field work is the fact that the response rate was different for different age groups, resulting in an over-representation of young and middle-aged people. But this is compensated by the fact that on week days younger- and middle-aged people are more often not at home. Figure 25 shows that the net result is that the sampled population was quite representative compared to the age distribution of the total rural population in the Marne. A chi-squared test to compare whether there is a significant difference between the two distributions resulted in a p-value of 0.72, indicating the null hypothesis that there is no significant difference between the two distributions cannot be rejected. Before conducting the chi-squared test, a Shapiro-Wilk normality test was carried out to control whether the data are normally distributed or not. The result was that the data were indeed normally distributed, both for the age of the respondents as for the age of all the citizens of the rural villages from existing databases. This means a chi-squared test could be carried out (Huang & Cui, 2015; Yap & Sim, 2011). Other minor sources of bias are people who may not tell the full truth and difficulties with understanding the local French dialect.

[5] In figure 25, people aged under 20 are left out because in the field work analysis only persons 18 years or older were considered an in the population data of existing datasets the data were only available for age groups, not for the separate ages and the 18-and 19-years old belong to the category 13-19 yr.

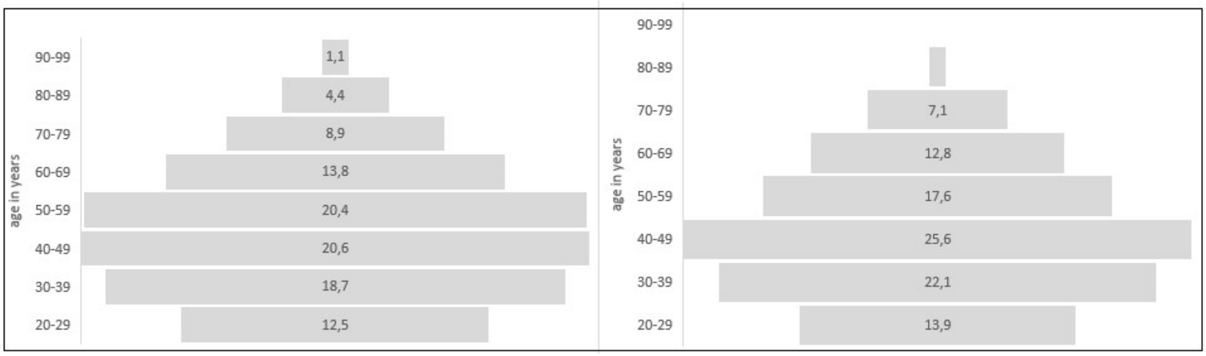


Figure 25: Population pyramids for (left) all the rural villages of the Marne department and (right) for the interviewed people. Own processing based on INSEE data for the left figure and own field work data for the right figure.

3.2.3. Data collection and data processing

All the information collected during the interviews was written down on the questionnaire papers. The data were consolidated by copying the answers into a structure spreadsheet. After this consolidation phase, the following data were available:

A. Data regarding population typology:

- Data on the age, education level, employment place, place of residence of the interviewed person and all the other persons living in the same house as the interviewed person. Mostly these include husbands/wife or children. For the children, who lived in the same house as the parents, the village where their school was located was also registered (if they still went to school).
- Information about the migration path of the interviewed person: where has he/she lived and for how long?
- Information about the age, education level, migration type, migration date of the children of the interviewed person. If the children of the interviewed person do not anymore in the same house, it was asked to where they were migrated, when and why they did so. Also the age and education level of the children was registered.
- Information about how much importance the person interviewed attaches to the services present in villages. For example, while living in a village, what's most important: the presence of school, a doctor or a minimarket?

The data were used to

(1) investigate which kind of migration types (rural to rural, rural to urban, urban to rural) are dominant
(2) analyze the relationship between different migration types and characteristics of the people such as age and education level
(3) examine which are according to the local population, the most important controlling factors of population evolution related to the number and type of services present in the village
(4) investigate the relationship between the importance people attach to possible services and the characteristics of the people like age and education. Those analyses not only provide useful information per se, but are also very important input data for the agent-based model that will be developed.

B. Data regarding typology of villages: Number and evolution of number of services. Service can be defined as all kind of elements that can be useful to be present in a village for living in that village such as the number of schools, shops, butcheries, bakeries, hairdressers, car repair shops and others. Those data were used to:

(1) investigate the relationships between absolute population numbers and absolute numbers of services present in a village
(2) investigate the relationship between population evolutions and the absolute number of services present in a village
(3) investigate the relationship between the population evolutions and the relative evolution of the number of services present in a village
(4) Link the number of services and the evolution of the number of services to the characteristics of the villages and the population living in the villages
5) Develop a typology of villages that can be used as an input in an agent-based model.

C. Other data: Additionally the following ancillary information was collected: number of dilapidated houses and degree of decline of a village, the number of recently built houses, information on the land use in the visited villages and, if the land use involves arable land, information about the crop type. In the case policy makers were interviewed background information on past policy decision and/or future policy plans was also registered.

3.1.3. Post-processing of raw data

Most of the collected variables on households and villages could be used straightforward. However, it was necessary to compile from the collected data four new synthesized variables: the Service Level (SL), Employment Potential (EP), utility level (UL) and Education Level (EL).

3.1.3.1 Service level in villages (SL)

The Service Level of a village (SL) is defined the weighted sum of services present in the villages and the distance to the nearest supermarket. The weighing makes it possible to incorporate the fact that some services are more important than others. The weights are based on the data obtained from field work where people were asked to indicate how important they find different possible services on a scale from 0 (totally not important) to 3 (very important). A weight for the different factors is calculated from this. An overview of the factors considered is listed in table 6. It is important to note that a result of the analysis is that people consider employment by far the most important factor, but this factor is synthesized separately (as employment potential (EP). Together with the EP, the SL contributes to the total utility level (UL) of a village. The weights of the SL are derived from interview data. For the SL of 2013, also the DTS was included. A multiple regression equation between the population evolution, the SL derived from field work and the DTS was used to determine the weight of the DTS within the SL. The SL is a significantly better predictor of the population evolution when the DTS is included. SL was only calculated for the villages visited during field work. For the other villages, the SL was extrapolated using the correlation between the SL on the one hand and population number and DTS on the other hand. The SL for the unvisited villages was assessed using the following equations, obtained from regression analysis.

$$SL\ 2013(village) = -0.14611 + population2013(village) * 0.0008471 + 0.0007668 * DTS \quad \text{(eq. 2)}$$

Since for the year 1982, there were no data on the location of the supermarkets in the Marne department, the SL for 1982 was calculated as follows

$$SL\ 1982(village) = -0.033 + population\ 1982\ (village) * 0.0008528 \quad \text{(eq. 3)}$$

Table 6: *Factors considered for calculating the utility level (UL)l related to services present in the village. The weights shown are for the whole population. Separate weights are also calculated for different age groups.*

Variable & description	*Weight 2013*	*Weight 1982*
Distance to the nearest supermarket (can be outside the village)	0.25	----
School factor (presence of nursery school and/or primary school in the village ; distance to nearest secondary school (can be outside the village))	0.20	0.27
Presence of minimarket in the village	*0.12*	*0.16*
Presence of bakery and/or butchery in the village	*0.12*	*0.16*
Presence of doctor and pharmacy in the village	*0.10*	*0.13*
Presence of café in the village	*0.08*	*0.11*
Public transport in the village and surroundings	*0.08*	*0.11*
Presence of sport clubs in the village	*0.05*	*0.06*

3.1.3.2. Calculation of the EP for a village

For each village, local employment data were available whereby local employment is defined as the number of people that is employed in the village. In order to include the employment possibilities in the wide surroundings for each village an EP was calculated following a procedure (equations 4 and 5) proposed by Poelmans et al (2010).

For practical reasons, it is assumed that for each village, all the employment is located in the center of the village. If there are no cities (sources of employment) within a reasonable commuting distance, this approach will result in that the local employment in the village itself is the most important factor in the resulting EP for that village. If there are cities not far away, than the city probably dominates in the resulting EP for the village, because the local employment is enormous higher in cities compared to rural villages. Boundary effects are taken into account for the calculation of the EP by adding employment data of settlements outside the department but nearby the border. The EP is calculated for different years: 1982, 1990, 1999, 2006 and 2013.

$$EP(X) = \frac{\sum w * El}{\sum w} \qquad \text{(eq. 4)}$$

$$w = \frac{1}{d^p} \quad \text{, p=1} \qquad \text{(eq. 5)}$$

3.1.3.3. Calculation of the utility level (UL)

SL and EP together determine the total Utility Level of a village (UL). Before calculating the UL, SL and EP were normalized. They are normalized in a way they both have a range between 0 and 1. In this study UL-values were calculated as weighted sums of normalized EP and SL, whereby the weights were assessed via model calibration (equation 6) – see chapter 4: Agent-based modelling.

$$UL = w1 * EP + w2 * SL \qquad \text{(eq. 6)}$$

3.1.3.4. Calculation of the education level

People were asked about their highest received education level. Based on their answers, they were divided into four categories: primary school, secondary school, higher education non-university, university. The variable is treated as a categorical variable.

3.1.3.5. Other variables

Some other used variables are straightforward and need no additional explanation. This is the case for the age of the people, the percentage young and old people, the population density and HP. Also variables related to the physical landscape are investigated, but their use is also straightforward and not requires additional information about methodology. All of the physical variables are described in the chapter ‘study area’.

3.1.4. Typology of population and typology of villages

Base on the post-processing the data collected during the field work and data from databases, a population typology and a village typology was developed. The population was divided into categories according age and highest received level of education. Next, the relationship between those different population categories and the occurrence of the different migration types (rural to rural, rural to urban, urban to rural, non-migrant) was investigated. A village typology was developed based on a cluster analysis using the village-related data described above. A maximum likelihood cluster procedure described by (Tan et al., 2005) was applied.

3.1.5 Correlation analysis

One of the aims of this study is to simulate and predict population evolution at the village level. The population evolution is therefore considered as a dependent variable while all other variables that were collected are independent variables. The bio-physical variables that were considered are soil fertility, land cover (forest fraction) and elevation. The socio-economic variables considered are among other local employment, the EP, the DTS, the SL of the village, the distance to the closes high school, the DTC and HP. All the variables are listed and described in Appendix B. Also information about the data source is listed in Appendix B. Also the intercorrelation of the investigated variables will be analyzed.

Two different data sets were used for the correlation analysis. The Marne department contains about 617 villages, approximately 580 of which are rural. The following two data sets are used for analysis:

- Dataset sample 1: the 24 villages visited during fieldwork
- Dataset sample 2: 40 Random selected rural villages

The second data set was used to check whether the findings of the correlation analysis based on the dataset consisting of the villages visited during field work are confirmed or not. If the results (r-squared values, significance values and equations) are close the same, it indicates that the correlation analysis based on the field work villages delivers representative results.

3.2. Results

3.2.1. Contextual data derived via fieldwork

This section aims to provide a context that can support the to understand the phenomenon of population dynamics on the basis of a qualitative description of the population evolution of a few representative village visited. All the visited villages are shown on figure 13 and to illustrate the village settings in the Marne department figure 26 shows a picture of the rural village of Congy with about 300 inhabitants.

Figure 26: The rural village of Congy. Picture taken on 15/07/2016 (own picture).

1 Examples of rural villages with a continuus decreasing population since 1968

12 out of the 24 selected villages are characterized by a predominantly declining population since 1968: Le Baizil, Binsont-Et-Orquigny, Bussy-Le-Chateau, Châtillon-sur-Marne, Coole, Congy, Corroy, Dommartin-Lettrée, Monthelon, Oyes, Venteuil and Pocancy.

For a long time, woodworking was the main source of local industrial employment in **Le Baizil.** Le Baizil is located in a forested area, so the land cover also played a role to the employment in this area. The employment in this sector decreased seriously in the ‘60s and ‘70s causing emigration of people and a decline of the population from about 350 to less than 250 inhabitants in about 10 years. After the fall in industrial employment, many services present in the village, like the school, the minimarket and the bakery closed. Le Baizil is an example of a rural village becoming partially abandoned due to a fall in industrial employment. Remarkable to Le Baizil is tha thte population dropped suddenly and was then followed by a period with a more constant population number (figure 27).

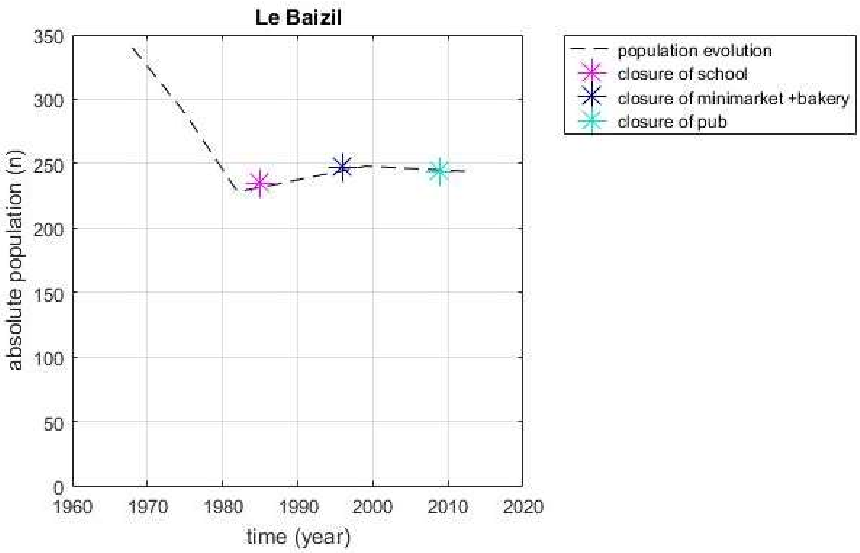


Figure 27: Le Baizil. The absolute population evolution over the years. Events (closure of services) in the village are indicated with an asterisk on the graph of the population evolution . The diameter of the asterisk signs indicates an uncertainty interval for the occurrence of the event. Own processing.

Congy is a village characterized by a strong rural abandonment. Here, the deterioration of buildings was the worst of all the rural villages visited, with many vacant and dilapidated houses (figure 28). In Congy, not far from Le Baizil, population decrease is not caused by fall of industrial employment, but by a fall in local agricultural employment: farmers abandoned the land. A comparison of the population graphs of Le Baizil and Congy makes clear that, however, both villages are characterized by a fall in population, the process of depopulation evolves quite different in Congy compared to Le Baizil. While in Le Baizil there was an abrupt population fall over a period of about 10 years followed by a stable population, the population has been continuously decreasing in Congy (figure 29).

An analysis of other population graphs of villages with declining population shows that Le Baizil and Congy are the examples of two prototypes: a type with sharp decline after the closure of an industrial activity and a type with a gradual but ongoing decline. The first type typically occurred in villages not too far from the urban centers, while the second type is present in the most remote locations where there has never been any off-farm activity. The villages were industrial closures were the main cause of abandonment are Le Baizil, Binson-Et-Orquigny. A decline in local agricultural employment was the main cause in Bussy-Le-Chateau, Coole, Congy, Corroy, Dommartin-Lettrée, Monthelon, Oyes and Ventueil. In **Pocancy,** there is a mix of both, namely a loss in agricultural employment and industrial closures that were equally contributing to the population decrease. This combined effect resulted in a population curve that combines a continuously slight to moderate decline and an accelerated population fall during the 90's after the industrial employment decreased (figure 30).

Businesses and services were mostly abandoned following an initial decline in the population. In Le Baizil, the depopulation triggered the school to close, followed by the minimarket and the bakery about ten years later and the closure of the pub another 10 years later (figure 27). Also the closure of the pub and minimarket in Congy occurred after the population fall (figure 29). The fact that a new hair dresser set up shop in a period of declining population (figure 29) illustrates the relation between the services present in a village and the population evolution is a complex process. However, tax exemptions for companies willing to settle in rural villages possibly played a role in the decision to open the hairdressing shop in Congy. Generally services decrease when population decrease, but the hairdresser example shows that exceptions are possible. In none of the other 11 visited villages with decreasing population did a new service open during a period with population fall.

Figure 28: Deterioration of Congy. Pictures taken on 15/07/2016 (own pictures).

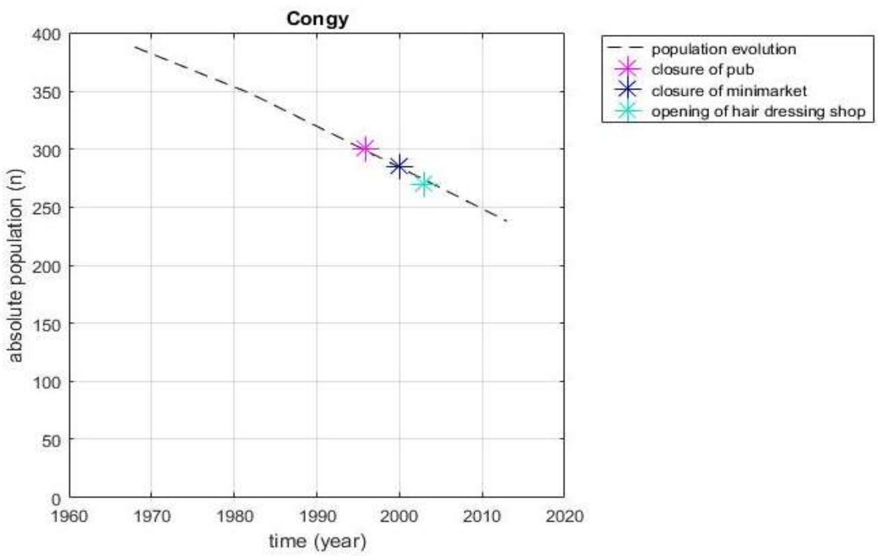


Figure 29: Congy. The absolute population evolution over the years. Events (closure of services) in the village are indicated with an asterisk on the graph of the population evolution. The diameter of the asterisk signs indicates an uncertainty interval for the occurrence of the event. Own processing.

In Pocancy (figure 30), the grocery store and restaurant closed already at the beginning of the period before the sharpest population fall. Before this considered period, the population was mostly constant for 20 years in Pocancy. This is possibly because the absolute population number is lower in Pocancy and was already too low for sustaining profitable services. A minimum number of villagers for most services seems to be 250 to 300 people. Also in Le Baizil and Congy the services started to close when the population number dropped below 250 people for Le Baizil and below 300 people for Congy.

After the last services are closed, the situation is often quite miserable for the people that stay. In certain villages, there aren't any services left (no minimarket, no bakery, no butcheries) and people have to travel up to 30km to reach the nearest shop. Figure 31 shows some pictures related to the closures of services. In some villages, the former services are partially substituted: in Dommartin-Lettrée, the bakery is closed, but the baker of a neighboring village tours every few days with his van in Dommartin-Lettrée, so the people of Dommartin-Lettrée can buy breads when he passes by (figure 32 left). In Congy, a bread vending machine replaces the bakery (figure 32 right).

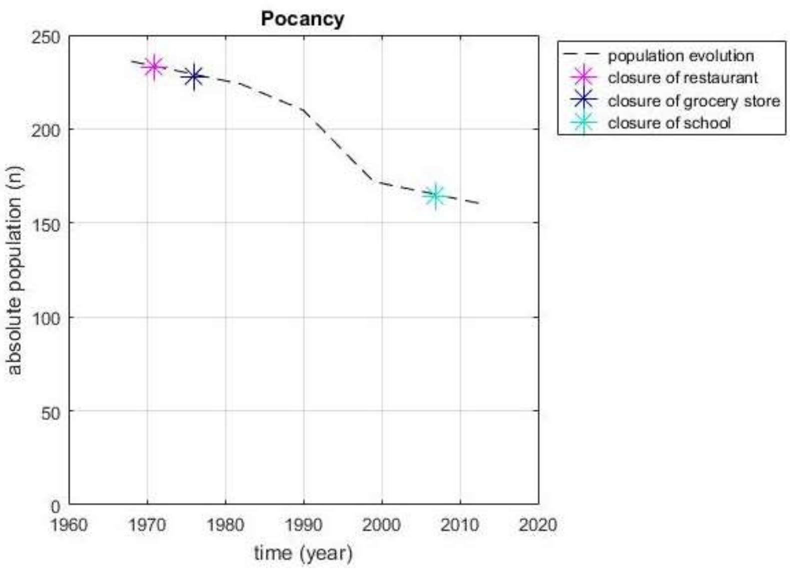


Figure 30: Pocancy. The absolute population evolution over the years. Events (closure of services) in the village are indicated with an asterisk on the graph of the population evolution. The diameter of the asterisk signs indicates an uncertainty interval for the occurence of the event. Own processing.

Figure 31: (a) The pub-restaurant closed a few years ago in Coole. Picture taken on 24/7/2016.(own picture). (b) former Fuel station. Picture taken on 16/07/2016 in Courjeonnet (own picture).

Figure 32: a) Moving baker from a neighboring village in Dommartin-Liettrée (left). Picture taken on 13/7/2016 (own picture) (b) bread automat in Congy (right) picture taken on 15/7/2016 (own picture).

<u>2 Examples of rural villages with a continues increasing population since 1968</u>

Nine out of the 24 selected villages are characterized by a predominantly increasing population since 1968: Brugny-Vaudancourt, Connantre, Ecury-Sur-Coole, Faux-Vésigneul, Jalons, Morangis, Pogny, Saint-Martin D'ablois and Vitry-La-Ville.

The cases of **Pogny** and **Connantre** will be described more in de tail to demonstrate that population increases in rural villages have various different causes. In Pogny, the population rise was initiated after the opening of a supermarket and commercial center in Pogny around 2000. The commercial center includes a supermarket, an optician, flower shop, sandwich bar and newspaper shop. This clearly was attracted people to move to Pogny (figure 33). In Connantre (figure 34), there is active industry ensuring employment for people in Connantre and the surrounding villages. A sugar factory and a stair factory are both located in Connantre and together they provide employment for hundreds of people (figure 35). In Pogny, the service sector is the dominant sector of employment with more than 60% of people employed in this sector. In Connantre, more than 60% of the people is employed in the industrial sector (calculated based on INSEE data).

When looking at the evolution of the number of services, it is noticeable that in Pogny it is the opening of new services which triggered an increase in the population, while in Connantre it is the other way around, an increase in population triggered new services to open a while after the population rise (compare figure 33 and 34).

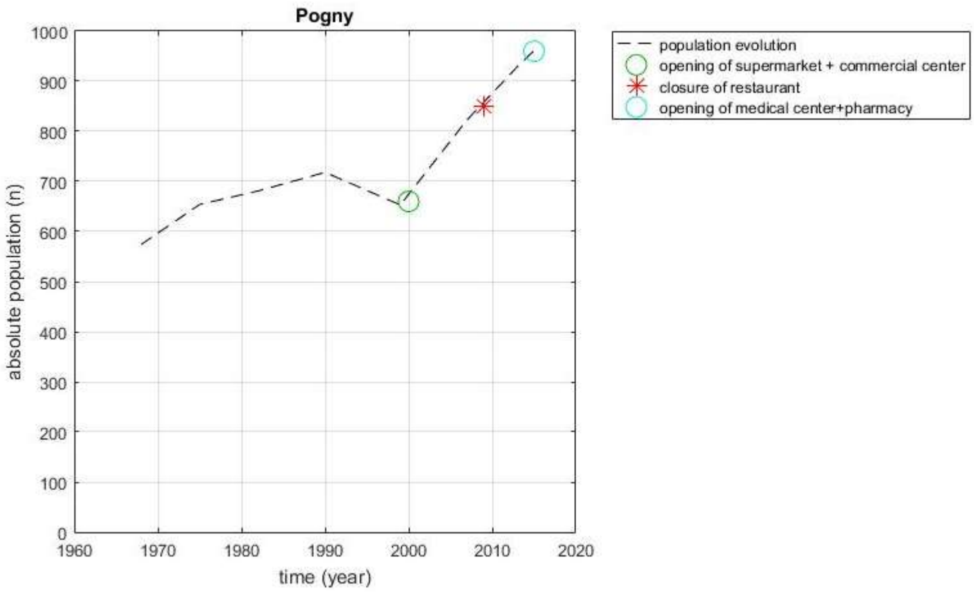


Figure 33: Pongy. The absolute population evolution over the years. Events (closures and opening of services) in the village are indicated . Events of opening are indicated with an open circle while events of closing are – as in the previous graphs – are indicated with an asterisk. The diameter of the asterisk signs and open circles indicate an uncertainty interval for the occurence of the event. Own processing.

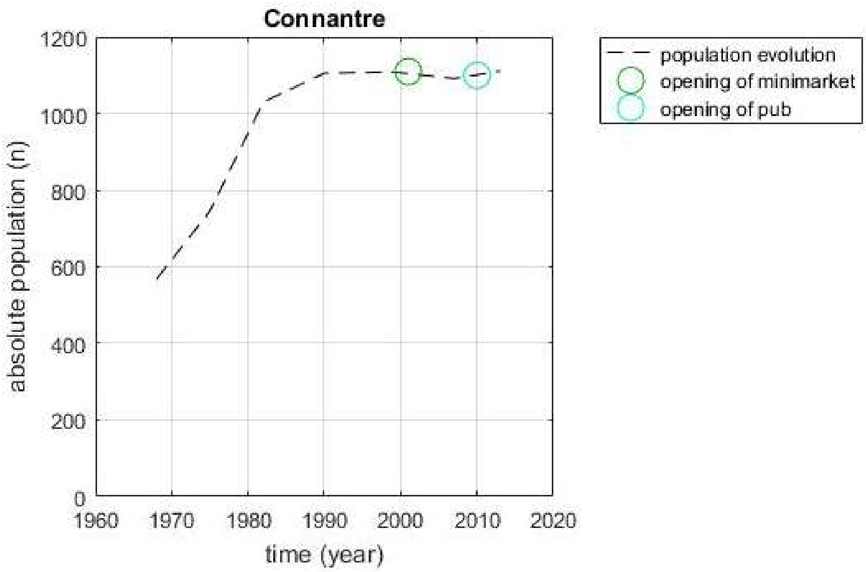


Figure 34: Connantre. The absolute population evolution over the years. Events in the village are indicated. The diameter of the open circles indicate an uncertainty interval for the occurrence of the event. purposefully made big enough so they cover an uncertainty interval for the occurrence of the event.

Figure 35: Active industry in Connantre. Pictures taken on 23/7/2016 (own pictures).

3 Examples of rural villages with about stable population since 1968

Two out of the 24 selected villages are characterized by a predominantly stable population number since 1968: Lachy and La Cheppe. Both have a population number around 300 and nearly no services. In Lachy, there is one nursery school located, in La Cheppe there are no services.

4 Examples of rural villages with inverting population trend during considered period (1968-2013)

Two out of the 24 selected villages are characterized by an inverting population trend during the period 1968-2013: Etoges and Châtillon-Sur-Marne.

Etoges is a special case. The population was decreasing and, services were closing in the 90's. It is worth noting that important services, such as schools and grocery stores start closing after the population falls below a threshold of around 250 to 300 people. Around the year 2000, the government parceled a large piece of land in upper-Etoges, resulting in the construction of around 100 new buildings (figure 36). This was attractive for young, recently graduated people because the land price for building was very low in Etoges. Some years after the government intervention, a new minimarket opened in Etoges. This is an example of the influence of policy interventions on population dynamics.

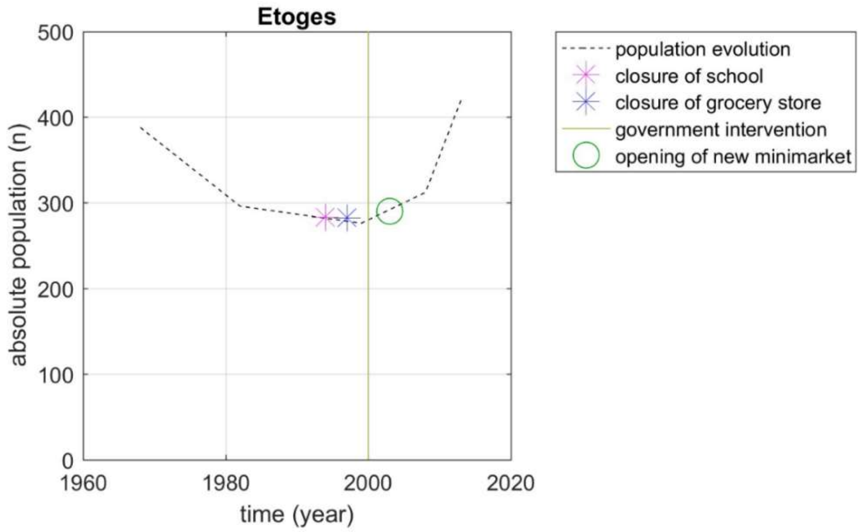


Figure 36: Etoges. The absolute population evolution over the years. Events (such as closures) in the village are indicated. The diameter of the asterisk signs indicate an uncertainty interval for the occurrence of the event.

Figure 37: New parceled district in Etoges. Picture taken on 17/7/2016 (own picture).

Another example of a policy intervention observed during field work is the shutdown of some railway connections. A few decades ago all train connections in the south of the Marne department were closed (figure 38). Since transportation networks are an important factor in population dynamics (see literature review), the lack of train connections has possibly influenced the population dynamics in this area of the Marne department.

Figure 38: Former railways in Connantre. Picture taken 23/07/2016 (own picture).

In **Châtillon-Sur-Marne** (figure 39), the population increased a little, until the grocery store closed. After which, there was a population decrease and more services close. Châtillon-sur-Marne is the only of the visited villages were the population decrease happened after the services started closing. In the other villages with decreasing population, the services always closed only after the population decrease.

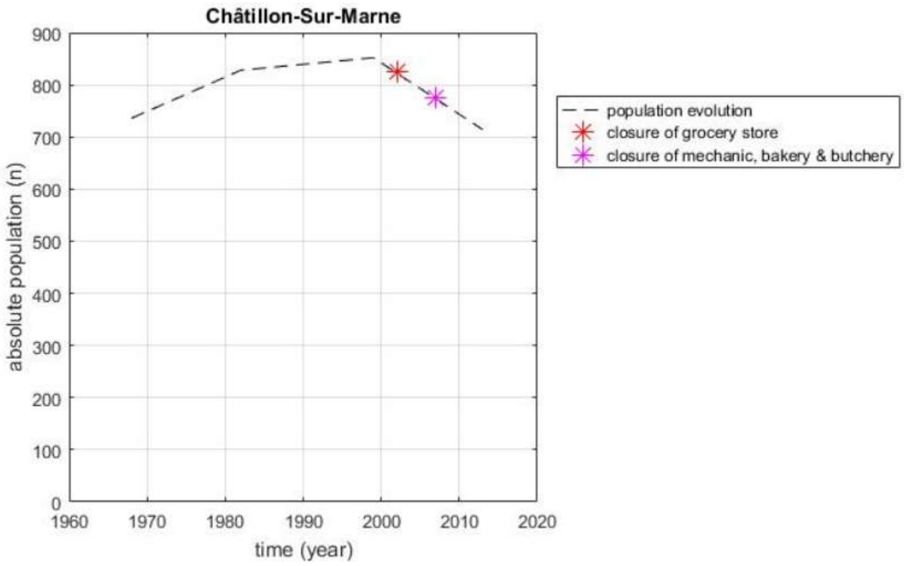


Figure 39: Châtillon-sur-Marne. The absolute population evolution over the years. Events in the village are indicated on the graph of the population evolution. Own processing.

From the qualitative analysis described above a few conclusions can be drawn:

Rural abandonment has mainly **2 different causes**: farmland abandonment and reduced agricultural employment on the one hand and industrial decline on the other hand. Both types of rural abandonment are characterized by a **different characteristic population evolution curves.** The fall in number of services in those villages happens mostly after the population decrease, although there are some rare exceptions where the closing of services triggers the population decrease.

Important services in a village like a nursery school, primary school, grocery store, bakery and so on have all the **tendency to close if the population drops below a certain threshold**. This threshold probably lies somewhere around 250 to 300 people.

For rural villages with an **increasing population**: this increase can have **various reasons**, like large employment sources in the village and surroundings (as can be observed in Connantre), but also the opening of a commercial center with a big number of service can initiate a rise in population like in Pogny or very low building land prices like in Étoges, because new pieces of land were recently parceled in this village.

There is **no unambiguous cause-effect relation** between opening/closing of services and the rise in population. An increase in the number of services can be both the cause as the effect of a population increase. However, the latter is more often the case.

Policy interventions can also play a role, in both directions. They can trigger both an increase and a decrease of the population.

3.2.2. Typology of population

It can be expected the type of migrations that people make depends on the characteristics of those people. Figure 40 and 41 show that age and education are valuable predictors regarding the type of migration people will decide to undertake. Education level is a better predictor for migration type than age. A categorical regression analysis – ANOVA (not shown) indicates that education level explains 45% of the variance of migration types, while the age only explains 17%. However, age and education level are also intercorrelated (figure 42) and combined they cannot significantly [6]better predict the occurrence of different migration types compared to education level. Of course, the fact that age and education combined can explain 45% of the variance (<multiple regression analysis) implies another 55% of the variance of the type of migrations is not explained by those variables. Possible other factors play a role such as personal preferences and income.

The main conclusions from the analysis are that better educated people are more mobile and migrate more often in comparison with lower educated people (figure 41). Rural-urban migrations occur predominantly with high educated people who hold an academic or higher education degree. People without such a degree only rarely make rural-urban migrations. In this group, rural-rural migrations are dominant. In the group with people that only went to primary school, urban -rural migrations frequently occur (figure 41).

[6] To avoid confusion, In this research is chosen to only use the word significant when it means there is a statistical significant relationship. In other contexts, the use of the word significant is avoided.

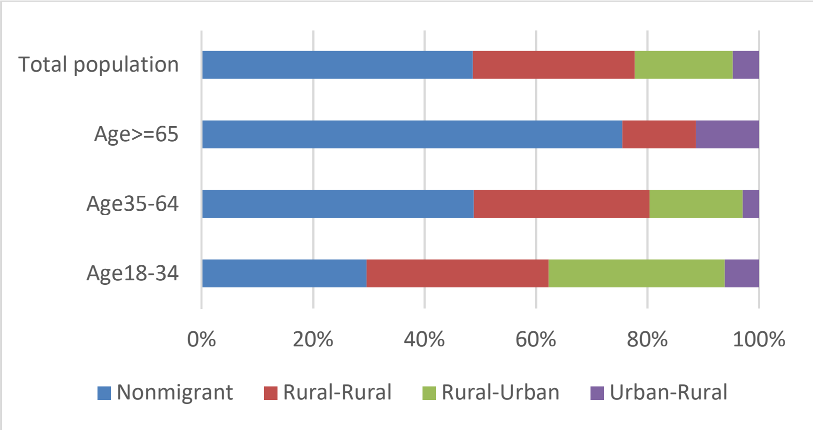


Figure 40: Occurrence of different types of migrations according to age group. Own processing.

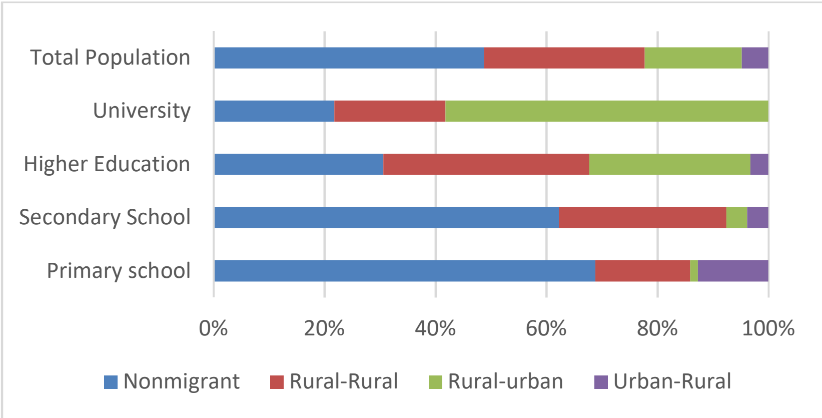


Figure 41: Occurrence of different types of migrations according to the highest received education level.

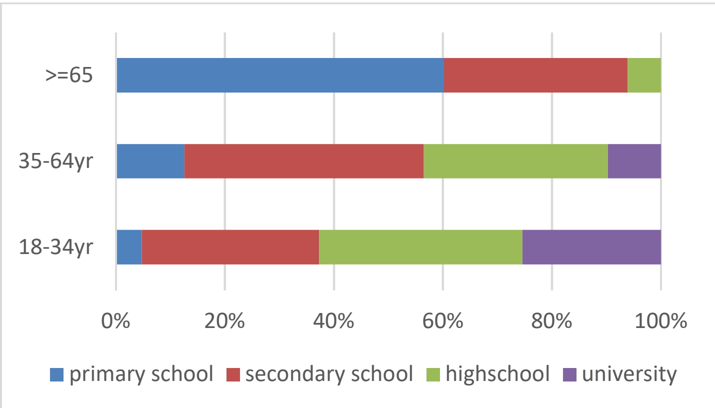


Figure 42: Highest received education level according to age. Own processing.

3.2.3. Typology of villages

The most important elements in a village that determine the willingness of people to live in a village or move to another village are (1) the available employment in the surroundings, (2) the number and type of services in a village and (3) the HP in a village, in that order. These findings could be translated into three quantifiable variables of a village: the EP of a village, the SL of a village and the HP in a village.

People who are graduated and not yet retired strongly prefer employment in the surroundings of the place of residence. The presence of services in the village are an additional beneficial factor. But this factor is mainly from crucial importance for older retired people, especially for the part of this group people that becomes less mobile and no longer able to drive a car or cycle. In that situation, a grocery store or minimarket in the village becomes a necessity. Third, HP are also important. HP does not trigger migrations, but once people have decided to move, they can play an important factor in determining what becomes the next place of residence. Poor people for example are in most cases not able to buy a house in an expensive area.

3.2.3.1. Employment Potential (EP)

Employment is high in the region around Reims, Châlons-en-Champagne and Epernay which are the most important sources of employment (figure 43). A low EP characterizes the east and south/southwest of the Marne department, where villages with a negative population evolution are situated (compare with figure 19 and figure 20). A decrease of the EP between 1982 and 2013 (figure 43) occurs usually in locations where the EP was already low in 1982. Especially villages where the employment was dominated by employment in the agricultural sector and to a lesser extent employment in the industrial sector were prone for a decline in employment (figure 43, figure 44). Figure 44 shows that the share of agricultural employment went down considerably in a lot of villages between 1982 and 2013. The share of agricultural employment to the total employment decreased in the majority of the villages while the share of employment in the services sector increased in the majority of villages (figure 44). The share of industrial employment decreased in about two third of the villages and increased in one third of the villages between 1982 and 2013 (figure 44).

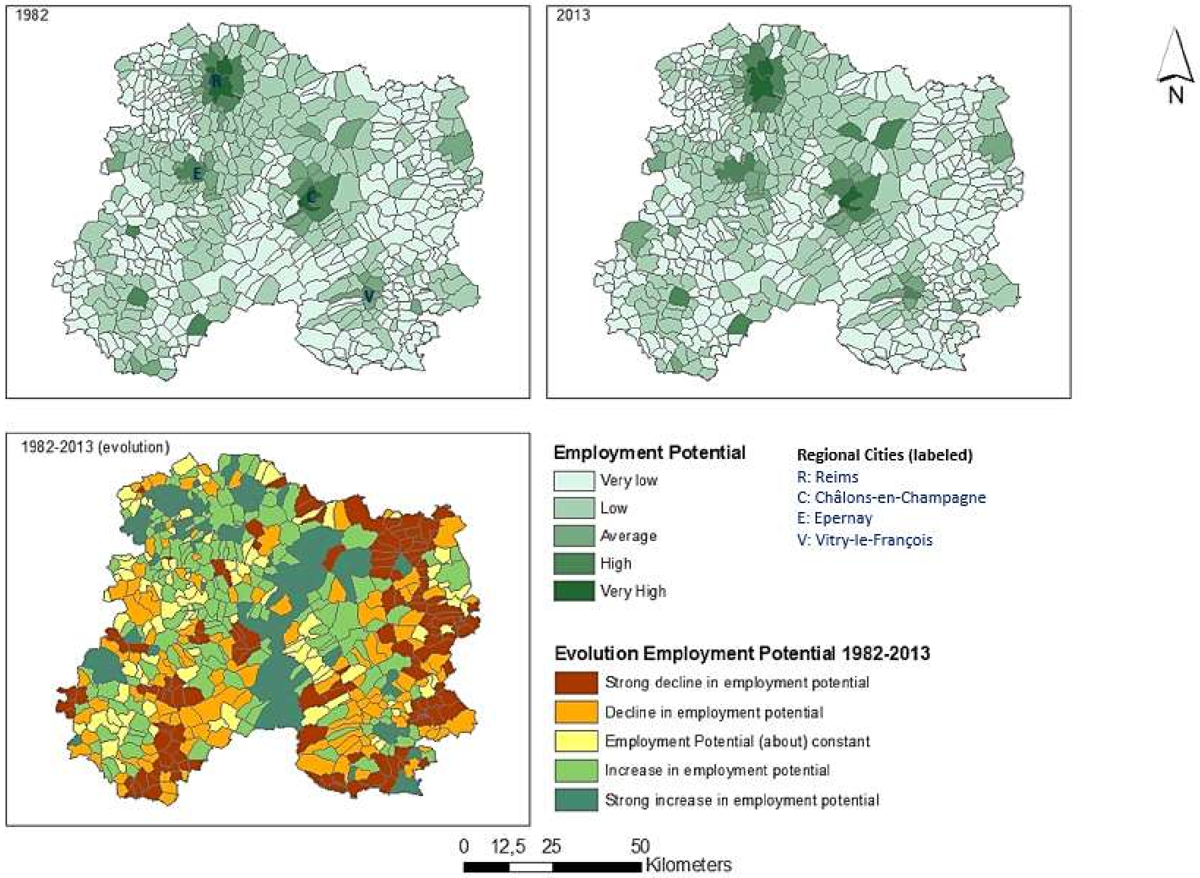


Figure 43: EP of the villages of the Marne department for 1982 (upper left) and 2013 (upper right). Also the evolution of the EP between 1982 and 2013 is visualized (bottom figure). Own processing.

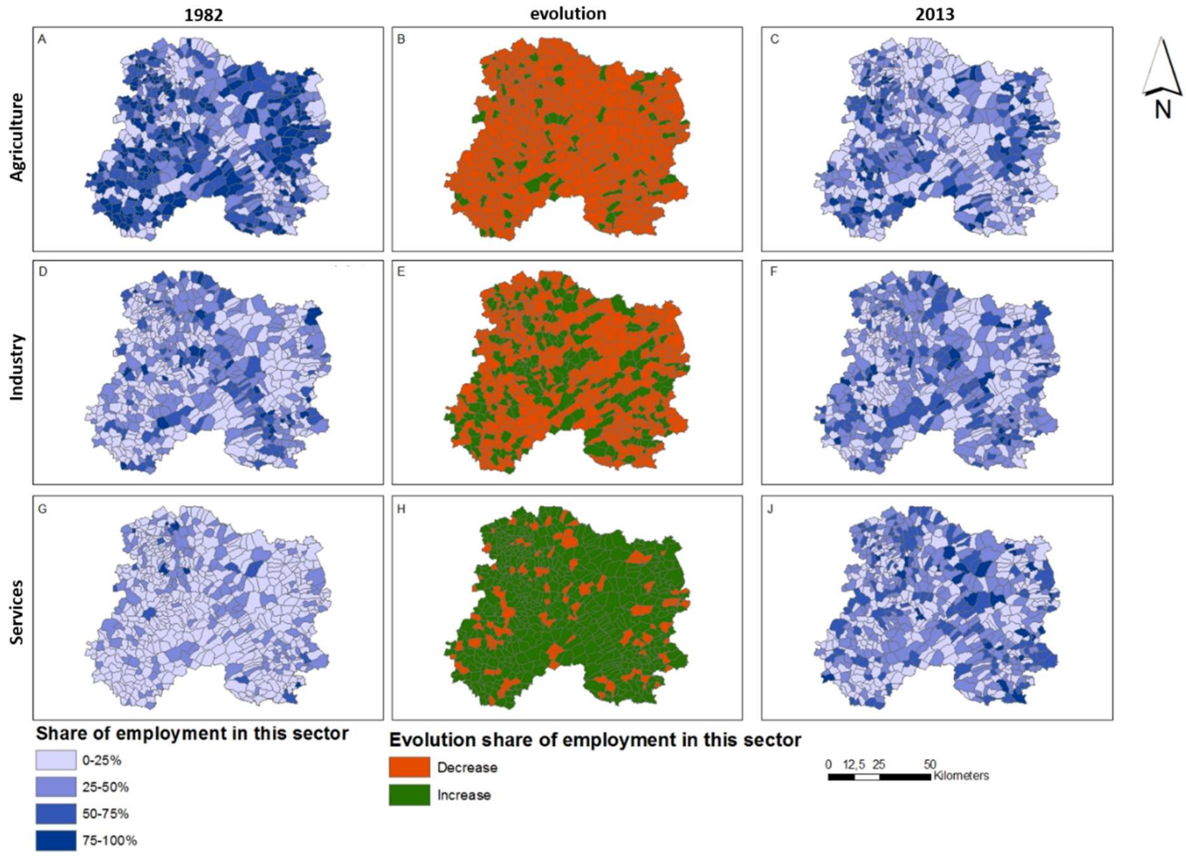


Figure 44: Evolution of the share of employment in the different sectors on the level of the individual villages. A) Share of employment in the agricultural sector in 1982 B) Evolution of share of employment in the agricultural sector between 1982 and 2013 C) Share of employment in the agricultural sector in 2013. Figures D-F provide the same information as A-C but for the industrial sector instead of the agricultural sector and figures G,H and J provide the same information for the services sector. Own processing.

3.2.3.2. Service Level (SL)

The distance of the supermarket is visualized in figure 45 for all the villages of the Marne department. The SL was only available for the 24 villages visited during field work and was for the other villages extrapolated based on regression equations as explained in the method section. The result is presented in figure 46. Because only rural villages were visited during field work, the extrapolation is limited to rural villages (figure 46). On the figure, it is clear that most rural villages in the Marne department have a low SL. In those villages, services are limited to a nursery school or a bakery or a bakery and a nursery school or in some cases no services were present. Some villages are better equipped, the spatial pattern of those villages is somewhat scattered, but they are nearly all located in the west of the Marne department.

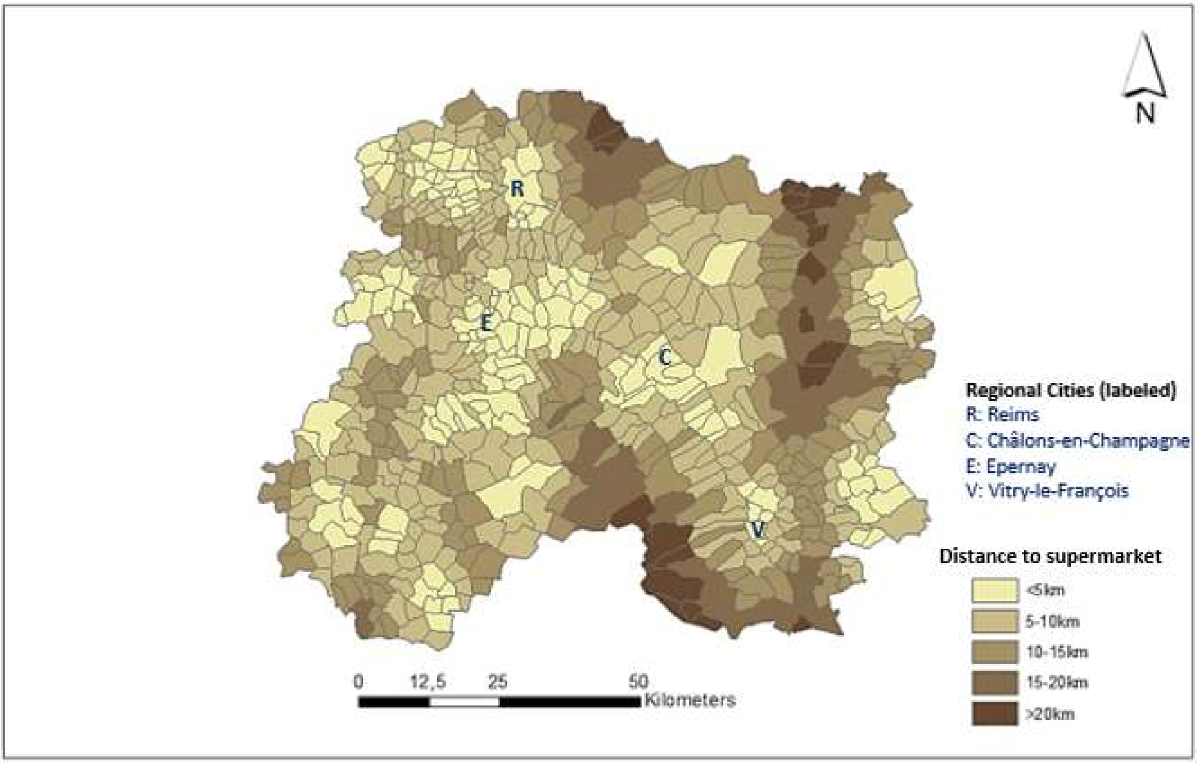


Figure 45: Distance to the nearest supermarket from the center of each village. Own processing.

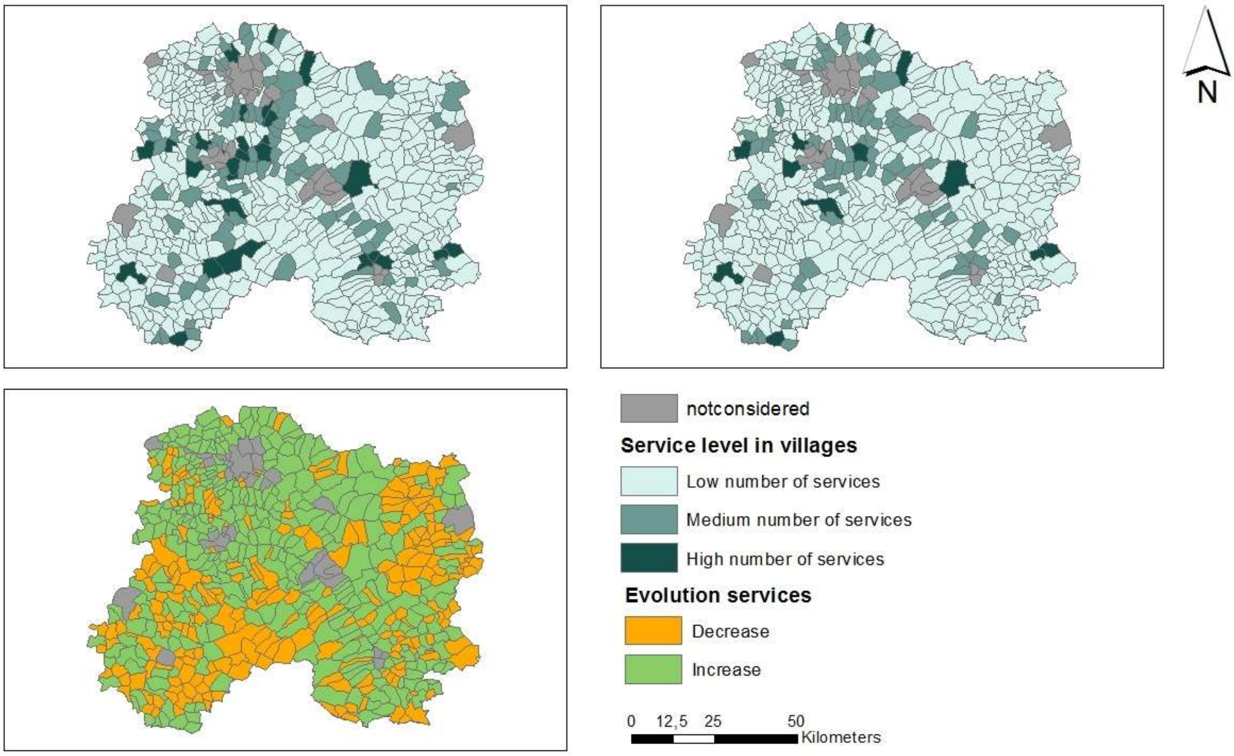


Figure 46: Service level in the rural villages of the Marne department extrapolated from field work data. SL is given for 1982 (upper left figure), 2016 (upper right figure) and the evolution between 1982 and 2016 (bottom figure). Own processing.

3.2.3.3. Housing Prices (HP)

The people living in rural villages in the Marne department often mentioned HP as an important factor for the selection of their migration destination. Various people mentioned they live in or moved to rural villages with few or no services because of the low HP. They can't afford to pay houses in the surroundings of thriving regions. However, there is a positive correlation between HP and the evolution of the population in a village (discussed in next section, 3.2.4. Correlation Analysis). Figure 47 provides the spatial pattern of the HP in the Marne department. The spatial pattern is more clustered than the services and the EP: very high HP around Reims, lower prices in the rest of the department. There are also similarities: the East and South/Southwest of the department have the lowest HP, coinciding with the regions with the lowest SL and lowest EP.

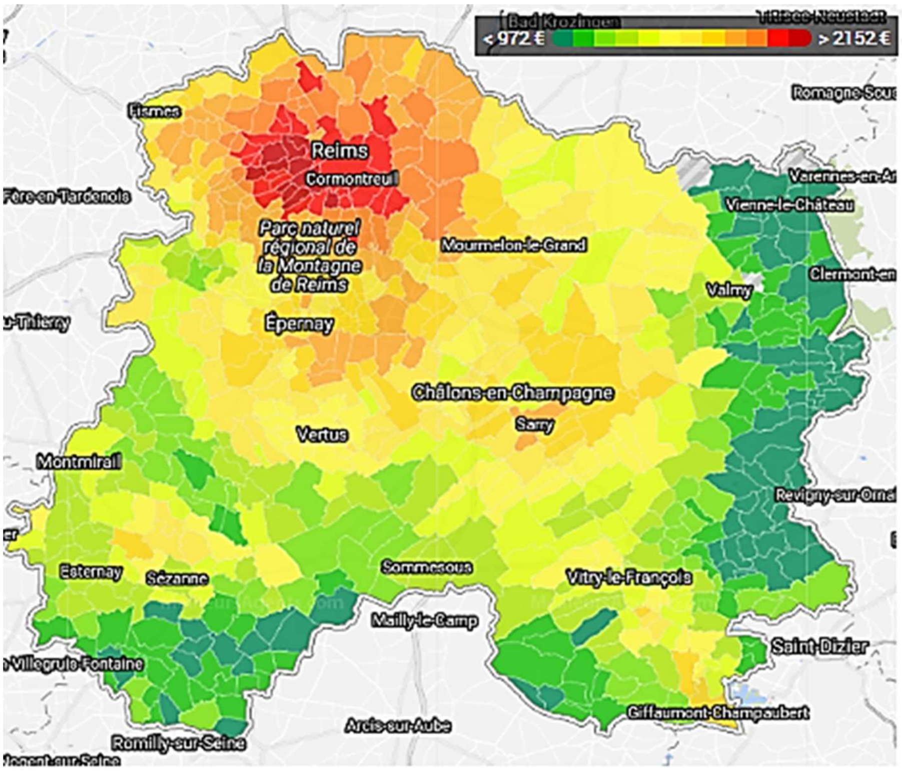


Figure 47: Housing prices November 2016. (MeilleursAgents,2016).

3.2.3.4. Link between typology of villages and typology of population

For the most important variable which explains the type of migrations, the education level, no data are available of the spatial pattern on the level of the villages. The spatial pattern of the fractions of young and old population are shown in figure 48. The spatial pattern of the young population cannot directly linked to previous discussed maps, as is the case for the spatial pattern of the middle-aged people and children (not shown). The spatial pattern of high fractions of old population coincide often with villages in the East and South/Southwest of the Marne department (figure 48). As discussed, this are the regions with the strongest decrease in rural population. This means ageing of the population is a consequence of rural depopulation.

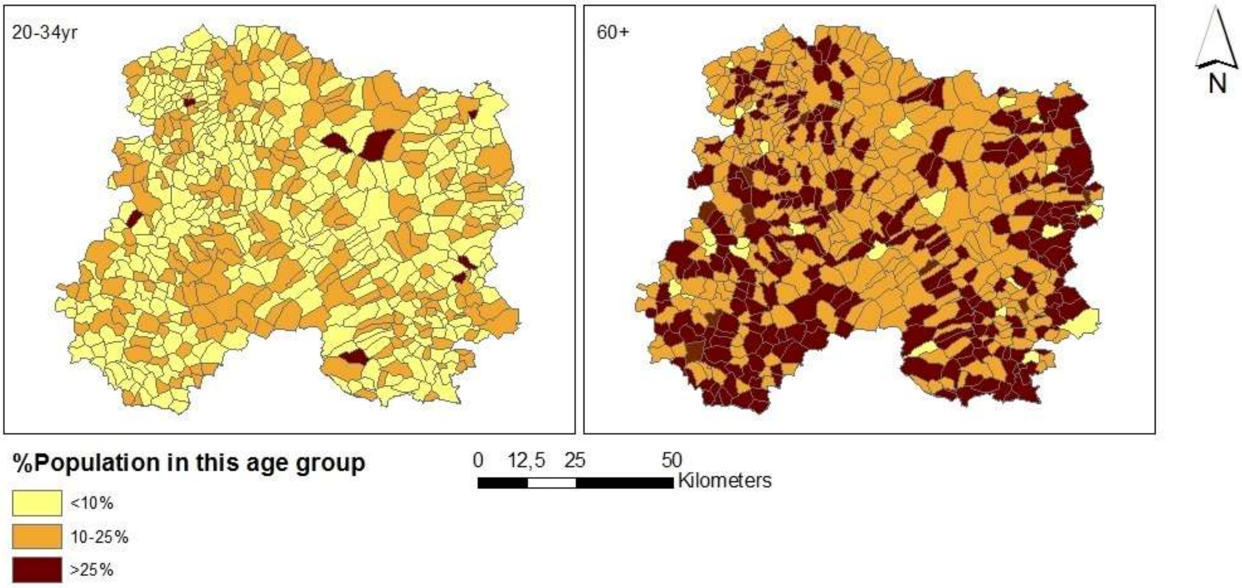


Figure 48: Spatial pattern of low and high fractions of young and old population. Own processing.

3.2.3.5. Cluster Analysis

Four clusters are presented, which comibned explain more than 85% of the variation of the population evolution are presented (figure 49).

In **cluster 1**, most of the villages have generally a decreasing population, but an increase of elderly population (60+). The EP is low and has a decreasing trend. The mean SL of the villages in this cluster is slightly above the average SL of the rural villages in the Marne department. In **cluster 2**, the villages have a decreasing population, and on average, the elderly population decreases as well, but to a lesser extent than the total population and the young population. The EP is low and has a slight decreasing trend. The average HP are very low for the villages in this cluster. The average SL for the villages in this cluster is around the same as the average SL for all the rural villages in the Marne department. In **cluster 3**, the characteristics of the villages are quite different from the characteristics of the villages in the two previous discussed clusters. On average, the population increases strongly for the villages in **cluster 3.** The elderly population stays almost constant for the rural villages in this cluster. The EP is high and shows an increasing trend. The mean of the HP and SL for the villages in this cluster is also high. **Cluster 4** is somewhat similar to cluster 2, but, for the most variables, the means of the variables are even more negative than in cluster 2. The EP and SL are truly low. The total population evolution is on average less negative for the rural villages in cluster 2. Both the young population and the elderly population are marked by a steeper decrease in cluster 4 than in cluster 2.

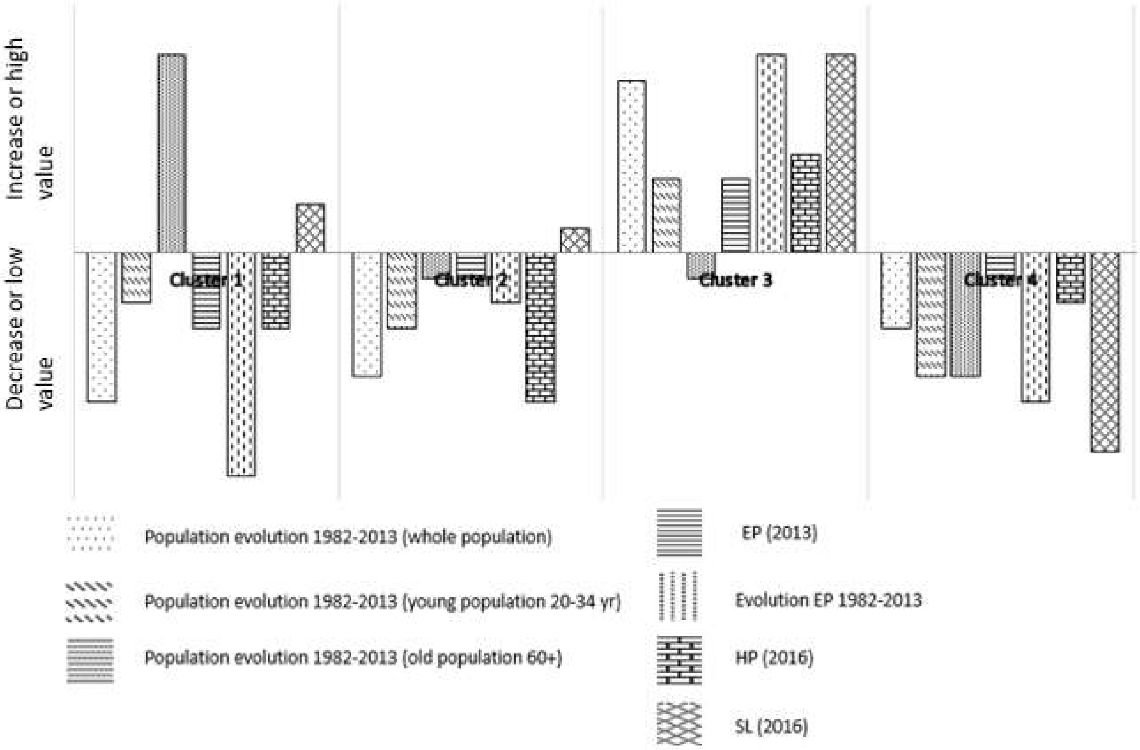


Figure 49: Cluster analysis typology of villages. Own processing using Rstudio software.

For each rural village, the similarity between the mean of each variable in the cluster analysis (for each cluster) and the value of that variable in the rural villages was calculated. Based on this, the rural villages are assigned to the category/cluster approximating their characteristics. The result (figure 50) shows the patterns detected earlier on, appear again. Cluster 4/Category 4, where the characteristics of the villages are the least favorable (very low EP and very low SL) is dominant in the northeastern region of the department while the majority of rural villages in the northwestern region and the center of the department is assigned to cluster 3 which has on average the most favorable village characteristics (high EP and SL).

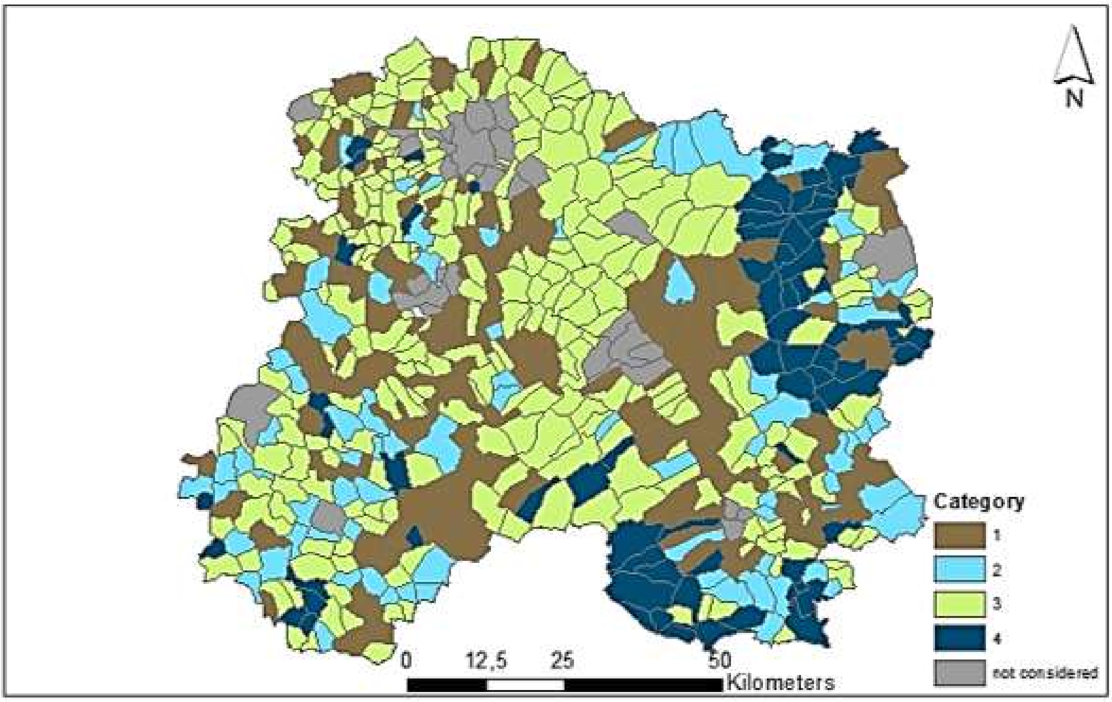


Figure 50: Classification of rural villages according to cluster analyses. Rural villages are assigned to the cluster for which its characteristics are the most similar to the mean characteristics of the cluster. Own processing.

3.2.4. Correlation Analysis

For the whole population, two data sets are used to check whether there is correlation between the relative population evolution (1982-2013) and the tested variables. The first data set consisted out of the 24 villages visited during field work, the second data set is a random test set containing 40 randomly selected villages. The purpose of the second set is to have an independent test whether the results of the correlation analysis based on the 24 visited villages during field work are reliable. The results are shown in table 7.

The results show a significant correlation between the relative population evolution between 1982 and 2013 (the dependent variable) and the following independent variables: (1) EP in 1982, (2) the evolution of the EP between 1982 and 2013, (3) the evolution of the number of services between 1982 and 2013, (4) the HP in 2016 and (5) the distance to the nearest supermarket (DTS) in a village (from the center of the village). The second test set confirms all those correlations. Additionally, the second dataset shows also significant correlations with the independent variables (1) distance to the nearest regional city (DTC) and (2) elevation.

As mentioned, the EP in 1982 is positively correlated with the population evolution, implying the EP from a certain year can be a useful indicator for the population evolution in the next years

Table 7: *Correlation values, significance values and correlation equations between the investigated independent variables and the dependent variable, the relative population evolution between 1982 and 2013. The analysis is for the whole population (all age groups combined). Own processing using raw data from INSEE and R-studio software for analyses. In the cases the correlations are market, there is a significant relationship.*

Variable	Correlation 1	p-value 1	Equation 1 Independent variable= relative population evolution 1982 until 2013	Correlation 2	P-value 2
%agricultural employment 1982	-0.04	0.49	NS	-0.09	0.06
%Service employment 1982	0.04	0.78	NS	0.04	0.13
EP1982	0.13	0.04	-0.19 + 0.44*EP1982	0.12	0.02
DEP 1982-2013	0.43	0.00	-0.01896 + 1.24145 *DEP	0.36	0.03
Number of services (normalized)	0.02	0.53	NS	No Data	
Evolution services (normalized)	0.48	0.00	0.13586 + 0.35904*DUT	No Data	
HP 2016	0.13	0.04	-1.05+ 0.0008*HP	0.09	0.03
DTS	0.20	0.01	0.376 -0.02323217*DTS	0.08	0.04
DTH	-0.02	0.44	NS	No Data	
DTC	0.04	0.17	NS	0.19	0.01
Elevation	-0.09	0.08	NS	-0.09	0.03
Fraction forest in villages	0.00	0.95	NS	-0.02	0.65
Soil fertility	-0.04	0.77	NS	No Data	
Population density	0.01	0.43	NS	0.06	0.08

But it is also interesting to take a look at the correlation between the population evolution and the variables for different groups of the population, namely between respectively young (20-34yr), middle aged (35-59yr) and old people (60+yr) and the population evolution. The results for all variables are listed in table 8. Some relevant correlations are also visualized. The relation between the EP and the population evolution is clearly stronger for younger people in comparison with older people. The yellow boxes on figure 51 indicate the area where the population evolution for the selected age group (left 20-34 years, right 60+ years) is negative and where the EP is equal to or lower than 0.8. It is visible that for an EP equal to or lower than 0.8 the population evolution is more negative for the younger population compared to the older population. Only 1 point with an EP </= 0.8 has a net positive population evolution for the young population, while there are a lot of villages with a EP <0.8 where the older population is increasing (figure 43)[7]. The strong correlation between the evolution of the number of services

[7] On the figures, the statistical terms R^2, confidence interval, prediction interval and p-value are used. If it is unclear what the exact meaning of those variables is or who those variables should be interpreted, more information regarding those elementary statistics is provided in Appendix C.

(normalized) and the population evolution is visualized in figure 52. The brown box on the figure indicates that the majority of the villages (19 out of 24) were the population evolution is negative, also have a negative evolution of the number of services in the villages.

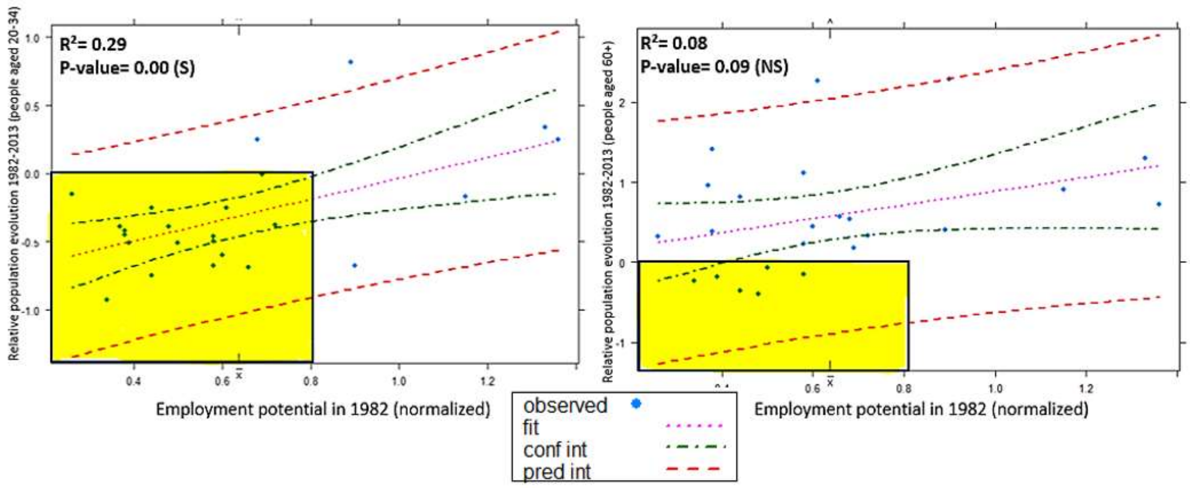


Figure 51: Correlation between the relative evolution of the population (1982-2013) and the EP for (left) the people aged between 20 and 34 years and (right) the people aged 60+ years. The yellow boxes indicate on both figures the area were the population evolution is negative and the EP in 1982 below or equal to 0.8. (S) is significant relationship and (NS)= not significant relationship. Own processing (using Rstudio software).

Table 8: *Correlation values, significance values and correlation equations between the investigated independent variables and the dependent variable, the relative population evolution between 1982 and 2013. The analysis is split into different age groups. Children (0-19 years) are not considered because they are assumed they aren't able or aren't allowed to take decisions independent. Own processing using raw data from Insee and R-studio software for analyses.*

Variable	Correl-ation 20-34yr	p-value 20-34 yr.	Equation 20-34yr.	Correlation 35-59 yr.	p-value 35-59 yr.	Equation 35-59 yr.	Correlation 60+yr.	p-value 60+ yr.	Equation 60+ yr.
%agricultural employment 1982	-0.04	0.70	NS	-0.04	0.60	NS	0.15	0.03	1.18 – 0.01 * AG82
%services employment 1982	0.02	0.43	NS	0.02	0.24	NS	0.01	0.26	NS
EP1982	0.29	0.00	-0.8024+ 0.7673*EP82	0.27	0.01	-0.8530+ 2.3869*EP82	0.08	0.09	NS
DEP 1982-2013	0.15	0.03	-0.39819+ 1.00818*DEP	0.27	0.01	0.3189+4.145*DEP	0.1918	0.02	0.4056+2.017*DEP
Normalized number of services 1982	-0.03	0.60	NS	-0.04	0.73	NS	0.00	0.30	NS
DSL 1982-2013	0.12	0.05	-0.278 + 0.25244*DSL	0.34	0.001	0.8441+ 1.2576*DSL	0.20	0.01	0.6544+0.5629* DUT
DTS	0.04	0.74	NS	0.00	0.34	NS	0.14	0.04	1.13818 – 0.04495* DTS
HP	0.30	0.00	-2.37 + 0.001513* HP	0.02	0.23	NS	-0.02	0.49	NS
Elevation	-0.07	0.10	NS	-0.06	0.13	NS	-0.32	0.02	2.016 - 0.0103*Elevation in m

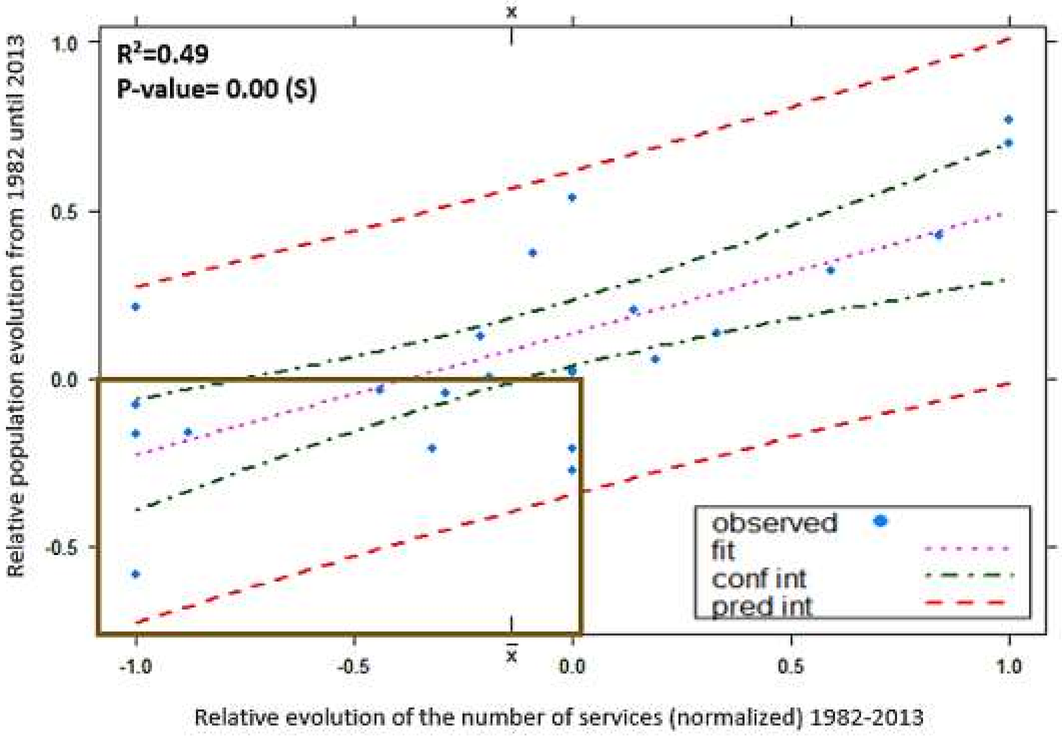


Figure 52: Correlation between the relative evolution of the population (1982-2013) and the relative evolution of the normalized number of services between 1982 and 2013. (S)= significant relationship and (NS)= not significant relationship. Own processing (using Rstudio software).

The relationship between the relative population evolution (1982-2013) and the HP is complex to understand. There is a net positive correlation between the HP and the population evolution (table 7, table 8, figure 53), but this is an example where the correlation analysis masks a part of the reality. On the one hand, the HP are a reflection of the state of the village: in villages that are thriving, where the EP is high and the EP is rising, where the SL is high, the HP will be higher compared to villages where there is no employment or there are no services, which is also clear from figure 54, which shows there is an important intercorrelation between the HP and the variables related to employment and the evolution of the number of services. That's why the positive correlation exists, seeing as people prefer the first type of village over the second type, certainly if the HP were equal. But, what clearly was a conclusion from the fieldwork conducted in the rural villages, people often mentioned 'low' HP as a factor to move to a village with few employment or little services. This implies the low HP acts as a negative feedback on the depopulation process. Following the same reasoning this also implies that the correlation between the HP and the population evolutions is much stronger for the younger population than for the middle-aged and older population (figure 45), because young people are more likely to move to a new place.

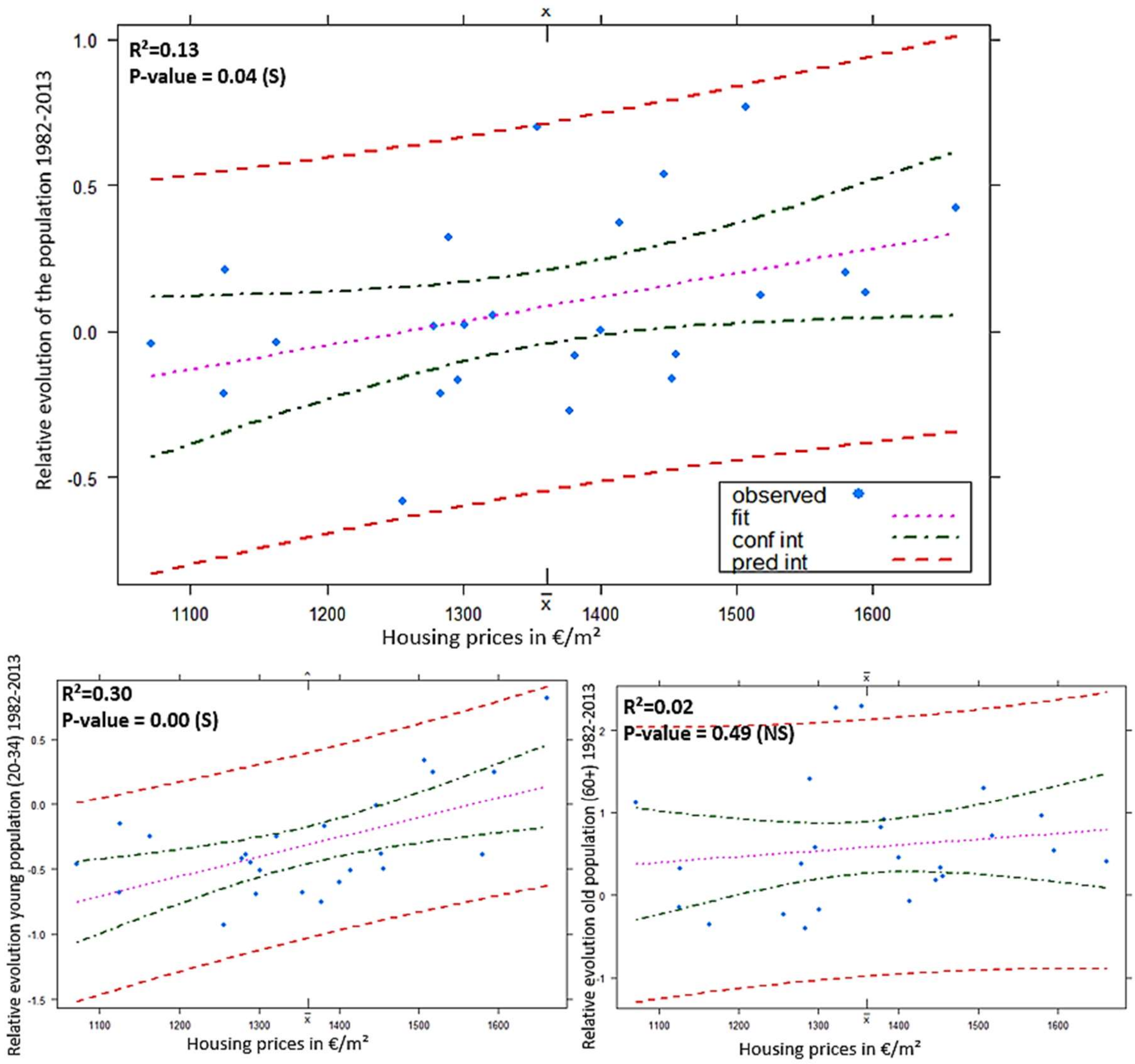


Figure 53: Correlation between the relative evolution of the population (1982-2013) and the HP for (upper figure) the whole population, (bottom figure left) the young people aged 20-34 years and (bottom right) the old people aged 60+. Own processing (using Rstudio software).

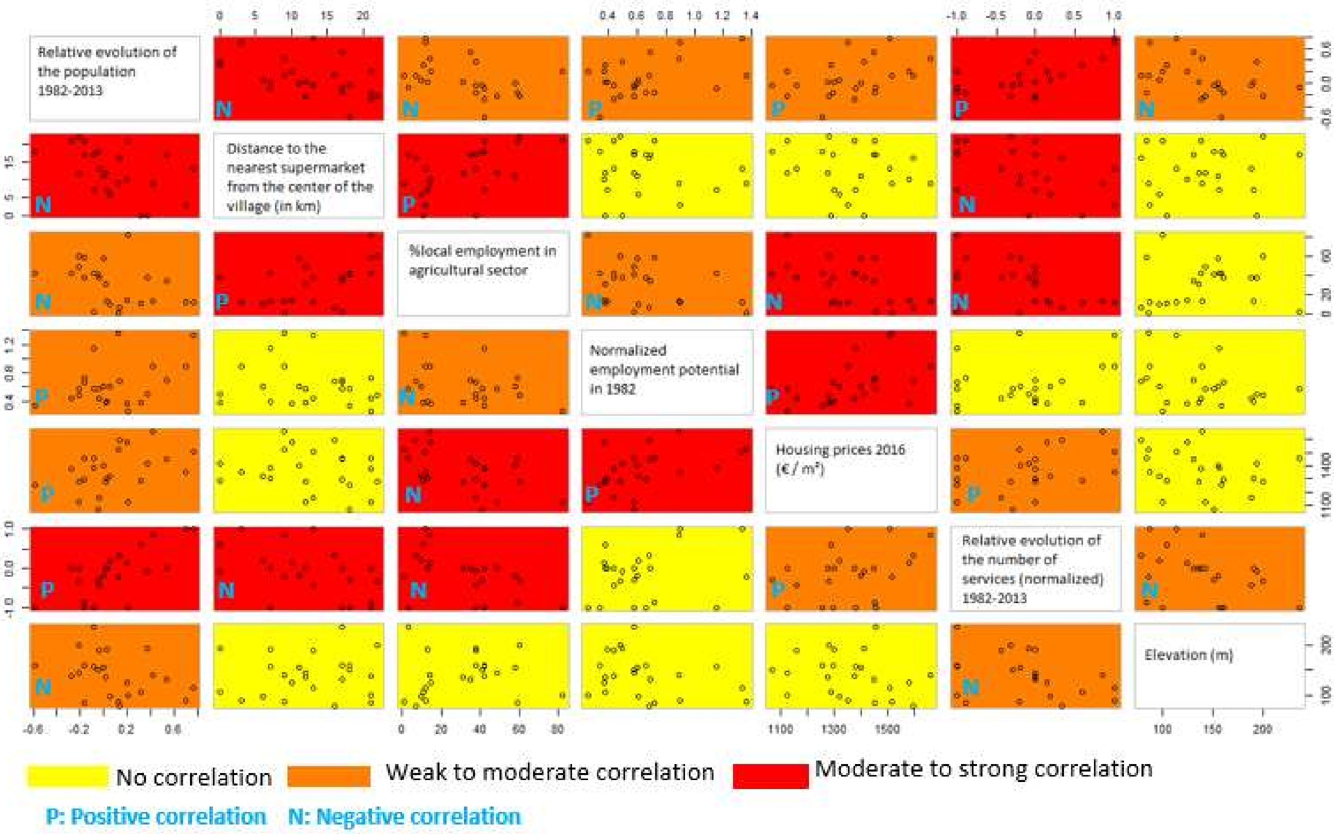


Figure 54: Scatterplots visualizing the correlation and intercorrelation between the tested variables. Colors indicate the strength of the correlation. Yellow= no correlation, orange= weak to moderate correlation and red= moderate to strong correlation. Own processing (using Rstudio software).

The intercorrelation between the variables (figure 54) is an important element in the detection of possible cause-consequence relationships between dependent and independent variables. The upper left variable is the relative evolution of the population, so the dependent variable. The figure shows all the independent variables that significantly correlate with the dependent variable. Intercorrelation was detected for all independent variables. All independent variables were intercorrelated with one or more other independent variables. For example, for the DTS from the center of the village, there is a moderate to strong intercorrelation with the fraction of local employment in the agricultural sector and the relative evolution of the number of services between 1982 and 2013.

CHAPTER 4: AGENT-BASED MODELLING OF RURAL POPULATION DYNAMICS

4.1. Data and methodology

A very comprehensive and detailed analysis of several factors influencing population dynamics in rural areas was carried out in the previous chapters. The final objective of this study is to develop a computational simulation model based on the knowledge gathered in the previous chapters of this research. It will be examined to what extent the collected data and the identified correlations can be used for population simulations. The goal of the model to simulate the population evolution for the period 1982-2013 which implies that simulations can be validated with the observed data. Finally population projections for the future can be simulated. Such simulations can serve policy makers by making clear what the impact of possible interventions on the rural land abandonment pattern can be.

Given the characteristics of the phenomenon to be simulated, i.e. decisions of individual people, the choice for agent-based model technique is designated. There are numerous examples that show that an agent-based model may have an added value compared to other simpler modelling approaches (Macal & North, 2010; Parker & Meretsky, 2004). One of the most important benefits is that in an agent-based model, individual decisions and heterogeneity can be included: not all people respond in the same way to external factors from the environment. An agent-based model integrates the factor that not all agents are the same (Parker & Meretsky, 2004) Additionally, the population can be split into groups (e.g. age groups) when simulating the population evolution . The agent-based model allows also for the inclusion of some randomness, which is a practical method to include for the fact different agents can take different decisions even when all external factors are the same. The agents in the model should migrate or stay, so the mobility of the agents trough space is an important consideration. The agents are the decision makers in the system . The decisions of the agents can be triggered by external factors from the environment like the SL and the EP. The agents react with their environment, the rural villages they live in. Those rural villages have a lot of characteristics that influence the population dynamics. Agent-based modelling also allows for the integration of dynamic interactions between the different characteristics of the villages (Schreinemachers & Berger, 2011). For example, the SL of a village will drop if the population falls below a certain threshold. In addition to the agent-based model, also an alternative model will be developed, namely a stepwise regression model (SRM), to evaluate whether an agent-based model has really an advantage in terms of performance in comparison with less complex models like regression models.

4.1.1. Development of an agent-based model

The population in the model was split into different age groups. The population is modelled separately for the following age groups: 0-19 years, 20-34 years, 35-59 years and 60+ years. Every year, the age of each agent increases with one year. As starting condition, it is assumed that the number of people are equal spread in the age groups 0-19, 20-34 and 35-59. This means that in the group 20-34 years for example, the amount of people 26 years old equals the amount of people who are 33 years, which equals the amount of people who are 25 years old and so on. For the last age group (60+ years), the exact age doesn't really matter, because people can't be move to the next age group in this case. And the number of deaths each year is based on real numbers: every year a fraction of the old people die, and this fraction is calculated based on real data from the Marne department.

During the development of a typology of the population and its relation to rural population dynamics, it also became clear that education is the most important factor in determining which type of migrations an agent will decide to make. Unfortunately, insufficient reliable data are available to include the education as a factor in the agent-based model. The EP and SL were used as main external factors influencing the population dynamics. The result of the previous analysis is that those are the most important factors along with the HP. The HP were only available for November 2016, no HP were found for 1982, so it was impossible to include reliable HP in the model.

4.1.1.1. Characteristics of the rural villages

For 1982, the starting year of the model, the real EP of 1982 is used. For future projections the real EP in the villages of 2013 can be used.

As previous analyses indicated, employment in the agricultural sector and employment in the industrial sector has been shrinking, while employment in the services sector has been rising. Based on the real EP in 1982 and the evolution of the share of employment in different sectors from 1982 to 2013 , equations for the yearly evolution of the EP are calculated from regression analysis. Because the share of employment in the agricultural sector and industrial sector shrinks and the share of employment in the service sector rises (figure 55), villages with a high share of agricultural or/and industrial employment are at higher risk to suffer from a decline of EP in the period 1982-2013.

The EP-equations derived from regression analysis were calibrated based on the real EP in the villages in 1990, 1999, 2006 and 2013 (eq. 8)

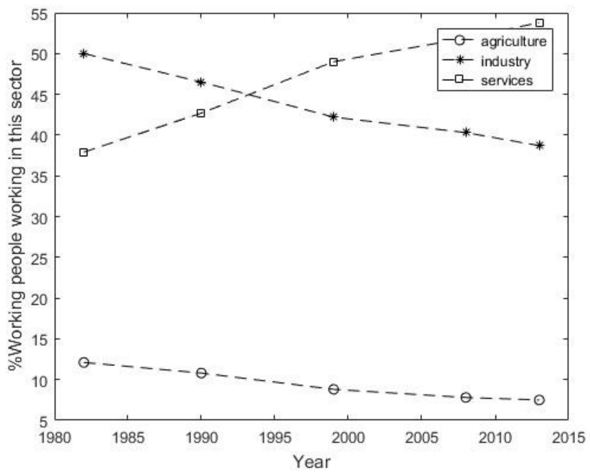


Figure 55: Evolution of the share of employment in the agricultural, industrial and services sector from 1982 until 2013 in the Marne department. Own processing.

SL: Available for 1982 and 2013 for each village (figure 46). In 1982, the real SL is used in the model. Evolution between 1982 and 2013 simulated and calibrated based on the real SL of 2013.

$$EP(yr, village) = EP(yr-1) * \%AG * 0.9985 + EP(year-1) * \%IND * 0.997 + EP(year-1) * \%SER * 1.0015 \quad \text{(eq. 7)}$$

With %AG= %employment in agricultural sector in the village and surroundings , %IND= %employment in industrial sector, %SERV= %employment in service sector , yr= year

$$EP(yr, village) = EP(yr-1) * \%AG * f1 + EP(year-1) * \%IND * f2 + EP(year-1) * \%SER * f3 \quad \text{(eq. 8)}$$

with f1,f2 and f3 are calibrated values for which the simulated EP is most similar to the observed EP in 1990, 1999, 2006 and 2013.

4.1.1.2. Migrations

A fraction of the agents residing in a village will decide to migrate to another village if the utility to stay (UTS) for the agent drops below a certain threshold. This UTS equals the EP for people aged 19-34 and 35-59 years. For people aged 60 years and older, the UTS equals the SL in a village. This means that for the younger and middle-aged people a decrease in EP triggers emigrations while for the older people a drop in the number of services can trigger an emigration from a village. This threshold where the UTS should fall below before the migration is initiated is determined based on the real total population evolution and the AMR for each age group. The threshold of the UTS corresponds around the median of the UTS of the villages. Above the threshold, people don't emigrate from the village. Below the threshold, the chance of emigrations increase. The lower the UTS, the higher the number of emigrations from the village. The Annual Migration Rate (AMR) for each age group was calculated based on field work data and literature. The AMR for each age group consists of two components:

AMRintra and AMRinter with AMRintra the inter-departmental migrations and AMRintra the intra-departmental migrations. Those migration rates are listed in table 9. People who migrate inter-departmentally disappear from the model. But, there are also people migrating from outside the Marne department to the Marne department. According to the field work data, the group of people leaving the Marne department is about five times larger compared to the group of people entering the Marne department. These people are nevertheless also taken into consideration (AMRinter in migration). They will be allocated to villages in the same way as the people who chose to make an intra-departmental migration.

Table 9: *AMR's for the Marne department calculated based on field work data. Own processing.*

	18-34years	**35-59years**	**60+years**	**Total population**
AMRinter(outmigr)	0.008	0.003	0.002	0.004
AMRinter(inmigration)	0.000	0.001	0.001	0.0008
AMRintra	0.023	0.017	0.008	0.017

Once an agent decides to migrate, they have to decide whether they migrate from the rural village to an urban area or to another rural village. Because only rural villages are considered in the model, the possible migration types are only rural-rural and rural-urban. For the occurrence of rural-rural and rural-urban migrations, data from field work were used. In the model, the type of migration depends only on age and not on education because lack of data about education.

If the agent migrates rural-urban, the agent disappears from the model. The agent will not be allocated to a new village but the sum of rural-urban migrations will be counted. If the agent chooses to make a rural-rural migration, the agent has to choose a **new village of residence**. It is more likely that the agent chooses a new village where the UL is high. But it happens sometimes that an agent decides to migrate to a village with a low UL.

To quantify this, the term probabilities is used. All the agents that have decided to migrate rural-rural are considered. For each of them, the probability that they will migrate rural-rural is 1, because in this phase we only consider the agents that finally have decided to migrate rural-rural. The probability that the agent moves to another rural village is 1. In his/her decision for a new place of residence, it is more likely the agent will migrate to a village were the UL is higher than the likelihood the agent will migrate to a village with a low UL, however, this is not impossible, but occurs less often. This is the concept illustrated on figure 56: The sum of probabilities is one – because the agent who has decide to move should choose another place of residence – and more residents will chose to move to villages with a high UL but a smaller fraction of residents will decide to move to a village where the UL is low (for example because a lower HP in such a village). How the UL is calculated is explained was the methodology section of chapter 3.

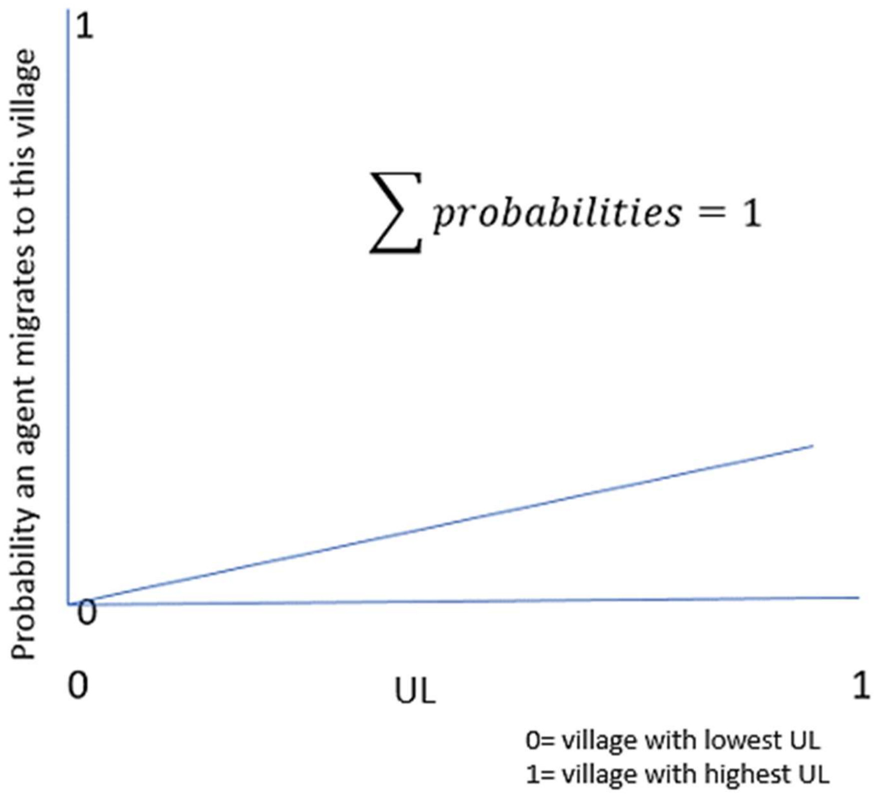


Figure 56: New destination probabilities. Once an agent has decided to migrate, there is a higher probability the agent moves to a location with a high UL (determined by EP and SL) compared to the probability he/she moves to a village where he/she has a low utility. Own processing.

Table 10 lists all parameters and variables used in the agent-based model. It also provides information about the source, confidence level and data type. The data type can be static (constant, the same over the period 1982-2013) or dynamic (updated every timestep). Figure 57 is a schematic presentation of the developed agent-based model.

Table 10: *Model input variables for the agent-based model with information about sources, use and confidence level of the used variables.*

Variable	Possible values	Data type	Data source	Comments	Confidence level
Fertility rate	Ratio (number of births per fertile women)	Dynamic attribute	Insee.fr	Fertility rate Marne department is used. Number of births depends on number of fertile women in the village.	High
Death Rate	Ratio (fraction of old people dying)	Dynamic attribute	Insee.fr	Death rate for the Marne department is used. It is assumed the Death rate is the same in each village (number of deaths depends on number of old people).	High
EP (normalized)	Between 0 and 1	Dynamic attribute	Insee.fr + Own analysis	based on raw data on local employment	Initial value: high From year 2: simulation
SL (normalized)	Between 0 and 1	Dynamic attribute	Field Work data	Weighted SL: if people indicated some services were more important than others, they received a higher weight in the calculation of the SL.	Medium
Population density	In inhabitants/km²	Dynamic attribute	Insee.fr	/	High
%employment agricultural sector	0 - 1 With 0=0% 1=100%	Dynamic attribute	Insee.fr	Data were available for 1982, 1990, 1999, 2006 and 2013. Based on this, 2 regression lines were calculated: one for the period 1982-1994 and one for 1995-2013. Period was split in 2 to make linear regression possible.	High
%employment industrial sector	0-1	Dynamic attribute	Insee.fr	Same method as %employment agricultural sector	High
%employment service sector	0-1	Dynamic attribute	Insee.fr	Same method as %employment agricultural sector	High
Migrations		Dynamic attribute	Field work data	Derived from field work data	Low too medium
Population number	Absolute number	Dynamic attribute	Insee.fr	/	High
Number of people <18yr	Absolute number	Dynamic attribute	Insee.fr	Yearly updated by the model based on simulated migrations and natural growth of the population	Initial value: high From year 2: simulation
Number of people 19-34 yr	Absolute number	Dynamic attribute	Insee.Fr	Yearly updated by the model based on simulated migrations and natural growth of the population	Initial value high From year 2: simulation
People 35-59yr	Absolute number	Dynamic attribute	Insee.Fr	Yearly updated by the model based on simulated migrations and natural growth of the population	Initial value high From year2: simulation
People 60+year	Absolute number	Dynamic attribute	Insee.Fr	Yearly updated by the model based on simulated migrations and natural growth of the population	Initial value high From year 2: simulation
Weights	Between 0 and 1 ;Sum of weights is 1	Static attribute	Field work& Multiple regression	/	Medium
Utility		Dynamic attribute	Field work	/	Medium
Utility treshold	Calibrated	Static attribute	Field work – calibration	/	low
Category calculation		Dynamic attribute			

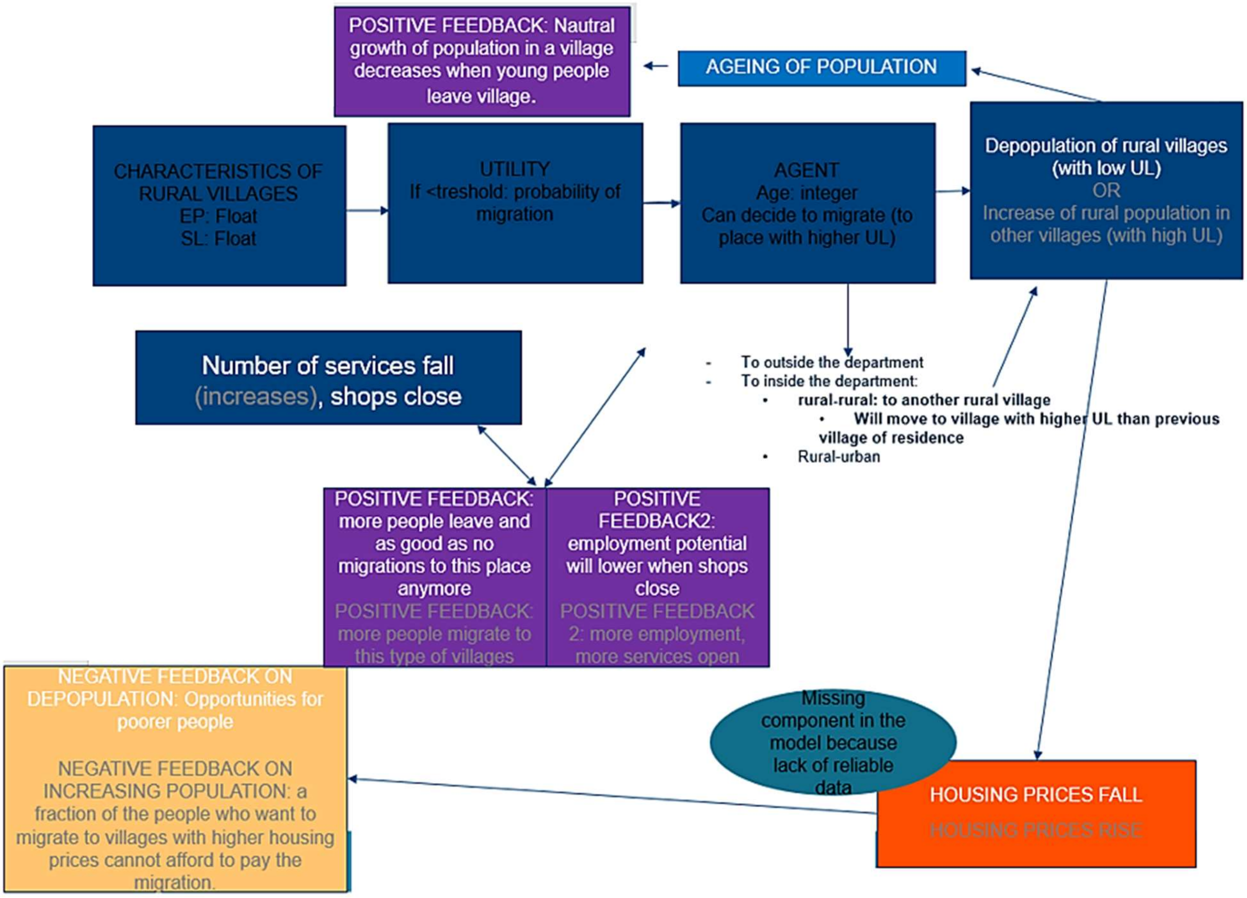


Figure 57: model components. Black text is applicable on all rural villages. White text is for rural villages experiencing depopulation and grey text is for rural villages experiencing increasing population. Black text is for all villages.

4.1.2. The Stepwise Regression Model (SRM)

The SRM is much less complex than the agent-based model. It only uses the output regression equations from a multiple regression analysis. This SRM is a multiple regression model for which the choice of the predictive variables is selected trough an automatic procedure (Sharma & Yu, 2015). The SRM selects the combination of predictive variables such that the resulting Aikaike information criterion (AIC) minimized. The AIC measures the relative quality of different statistical models, in this case multiple regression models, for a certain dataset. The SRM starts from a full regression model containing all possible predictive variables. Iteratively, more predictive variables are eliminated and the final selected model is the model with the lowest AIC (Zhang, 2016). The AIC tells nothing about the absolute performance of a model, but is a handy tool comparing the relative performance of different models. The calculation of the AIC value is based on the maximum likelihood function and the number of parameters used. Incorporating the last in the calculation helps avoid overfitting.

4.1.3. Model Calibration and Model Validation

For the population evolution on the level of the individual villages, various error measures will be used for both model calibration and model validation. For the validation different datasets will be used. The model calibration datasets consists of 100 rural villages including all the 24 villages visited during field work. Those villages are included because the input data of those villages are real data and none of the data are based on extrapolation. However 24 villages might be a too small sample for model calibration, therefore additionally 76 randomly selected villages are added. The model validation dataset consist of a random selected sample of 100 rural villages. Villages used for model calibration are excluded from this analysis.

The error measures that will be used are the Root Mean Square Error (RMSE), the Nash-Sutcliff Efficiency (NSE) and the coefficient of determination (R^2). The RMSE is a widely used error statistic in literature used in various domains and is often considered as one of the best error measures. They are more appropriate and a more stable statistic assessing model performance than for example the mean absolute error (Chai & Draxler, 2014; Mayer & Butler, 1993). RMSE is used under the assumption that the errors have a normal distribution (Chai & Draxler, 2014). Disadvantages of the RMSE are that the error measure is absolute and scale dependent. The error depends on the number of people living in a rural village. Villages with a higher population will have a higher RMSE most of the time (Hyndman & Koehler, 2005). A second disadvantage is that outliers can heavily influence the result. This error measure can be used to assess the progress of model performance during model calibration, but is not recommended to use when comparing two different datasets, implying this error measure cannot be used for model validation (Shcherbakov et al., 2013). Another error measure is the Nash Sutcliff Efficiency (NSE). The NSE was initially developed to evaluate models in hydrology, but can also be suitable for evaluating models in other domains (Moriasi et al., 2007). The advantage of the NSE over the RMSE is that the calculated number provides information about the absolute performance of a model. A NSE of 0 indicates that the observed mean can as good predict the population evolution as the developed model (Moriasi et al., 2007). The R-squared value offers this very same advantage, where a value of zero indicates the model performance is not superior to the performance of a random model.

The model calibration procedure has two steps. First, the errors will be minimized by calibrating the EP. In the model, the evolution of the employment model (since 1982) was based on the share and evolution of the share of agricultural, industrial and service employment in each village. Based on the real data of the EP in 1990, 1999, 2006 and 2013, the modelled EP will be calibrated. In the second step, the model will be calibrated by adjusting the weights of the factors (EP and SL) that are used to allocate people to a new village when they have decided to migrate. An automatic calibration algorithm is developed using Matlab software. The algorithm automatically tests the performance of the model (using the mentioned error measures) for a range of values of the weights of the EP and SL. It is investigated for which combination of the two weights the error is the lowest. The best model performance occurs were the RMSE is minimum and the NSE is maximum. In a first step, the algorithm searches the minimum error over a large range of weights and iteratively there is zoomed in on better areas.

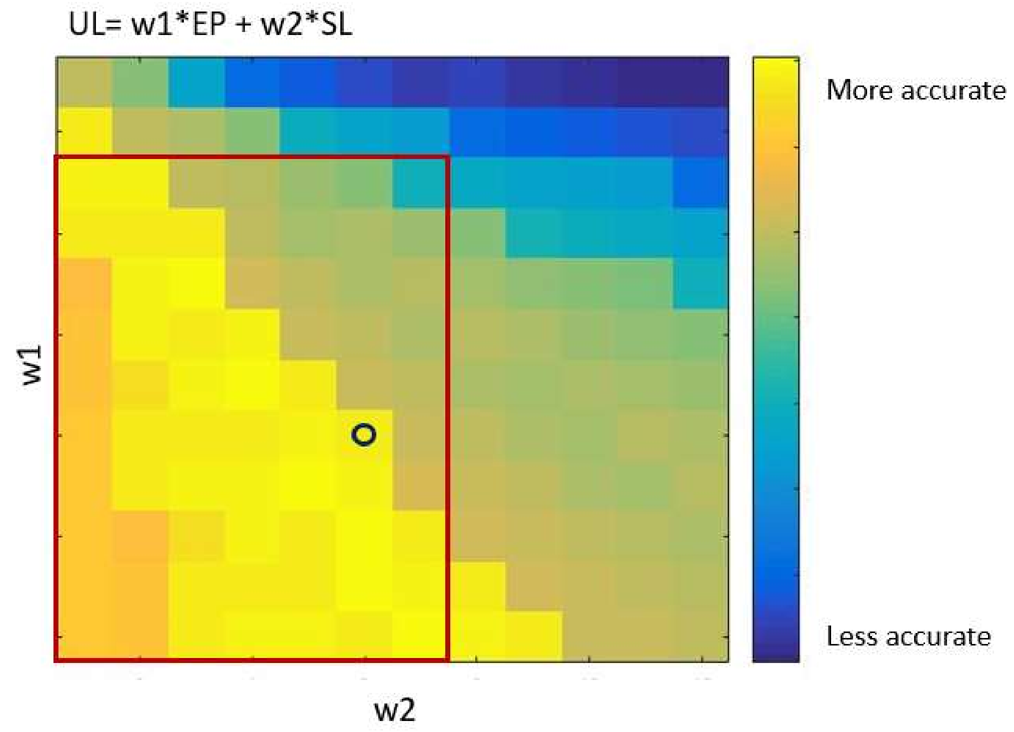


Figure 58: illustration of the modem calibration procedure. Suppose the open black circle indicates the initial weights derived from data like field work data (this is an illustration of the methodology, the example on the figure are not the real numbers). At a first calibrations step, the error is calculated for a large range of possible values for w1 and w2. In a second step, a domain within the initial domain is selected where the errors are lower (the values more accurate). On the shown figure for example, this can be the area indicated with the dark red box. The procedure is repeated (further zoom-in in each time-step) until the spatial variation in error is disappeared. Own processing.

Based on the discussed error measures, the performance of the calibrated agent-based model will be compared with the performance of the uncalibrated agent-based model, the SRM, a random model and a null model. In a random model, the global population evolution is correct, but the spatial allocation of people is random and not dependent on the controlling factors . A null model is a model that predicts no change compared to the initial situation and predicts the population will be constant in all villages during the simulated period (Poelmans & Van Rompaey, 2010).

Not for all years real data about population and employment are available. Therefore, the model is calibrated and validated based on some selected years for which all data are available: 1990, 1999, 2007 and 2013. Both calibration and validation are executed both for the total population as for each age group separately, except the age group of the children (0-19 year) because they are assumed to not be able or allowed to migrate independently. As mentioned, for calibration and validation different datasets are used. Table 11 lists the error measures that are used.

Table 11: *Used error measures, their formulas, possible values and interpretation.*

Error measure	Possible values	Interpretation
RMS $RMS = \sqrt{\frac{\sum(\hat{y}_i - y_i)^2}{n}}$ (eq. 9) $\hat{y}_i = observed; y_i = model\ prediction$	0 to +∞	Difficult to interpret because value depends on values of dataset, scale dependent In general: how lower, how better Possible to compare 2 different models if same data set is used For comparing the results before and after calibrating the model, it is more difficult to use because for model calibration and validation different datasets are required.
Nash-Sutcliffe $NSE = 1 - \frac{\sum(\bar{y}_i - y_i)^2}{\sum(\widehat{y}_i - mean(\hat{y}_i))}$ (eq. 10)	-∞ to +1	If NSE < 0, the observed mean is a superior predictor compared to the model If NSE = 1, the model is perfect A bit similar to NRMS
R²-error $R^2 = 1 - \frac{\sum(\bar{y}_i - y_i)^2}{var(\hat{y}_i)}$ (eq. 11)	-1 to 1	1= perfect model 0= random model

Also binary maps (increase/decrease of population) are produced, modelling the population evolution in all villages of the Marne department. Confusion matrices are used to validate those maps. Those matrices are a suitable way for evaluating binary classification maps (Sokolova & Lapalme, 2009). The accuracy of the model outputs (% correct classified villages) and the Kappa Index of Agreement are calculated. The Kappa agreement corrects for the fact that the model has predicted the direction of the population evolution correct by chance (Viera & Garrett, 2005).

Table 12: *Confusion matrix.*

		Predicted	
		Decrease	Increase
Observed	Decrease	A	B
	Increase	C	D

$$Accuracy = \frac{A+D}{B+C} \qquad \text{(eq. 12)}$$

$$Kappa = \frac{Accuracy \quad (PYC+PNC)}{1-(PYC+P \quad)} \qquad \text{(eq. 13)}$$

$$PYC = Probability\ of\ yes\ by\ chance\ \ ; PNC = Probability\ of\ no\ by\ chance$$

$$PYC = \frac{A+B}{A+B+C+D} * \frac{A+C}{A+B+C+D} \qquad \text{(eq. 14)}$$

$$PNC = \frac{C+D}{A+B+C+D} * \frac{B+D}{A+B+C+D} \qquad \text{(eq. 15)}$$

4.2. Results

4.2.1. Model Performance

According to all error measures, both the uncalibrated agent based model (ABI) and calibrated agent based model (ABC) perform better than a stepwise regression model. When the ABC is applied on an dataset not used for model calibration (validated model- ABV), this validation shows the calibration was successful in increasing model performance (table 13). In terms of absolute performance, the model performs better compared a model using the absolute mean to predict population evolutions because the NSE>0 (table 13). However in the last year tested, 2013, the model simulations of the ABV become worse than simulating the observed mean for the whole population. However, for each age group separately, the NSE remains also in 2013 above zero (table 14). In a comparative analyses, the ABV also clearly performs superior compared to more simple models like a SRM (table 13).

Table 13: *Comparison of performance of the uncalibrated initial agent-based model (ABI) and the calibrated agent based model with a stepwise Regression model (SRM), a random model (RM) and a model that models the observed mean of the data (OM). Performance is measured trough the Nash-Sutcliffe Efficiency (NSE), R² value and Root Mean Square Error (RMSE).*

Model type	*NSE*	*R²*	*RMSE*
ABI	*-0.45*	*0.89*	*74.02*
ABC	*-0.08*	*0.90*	*42.21*
ABV	*-0.34*	*0.89*	*53.38*
SRM	*-0.69*	*0.63*	*178.29*
RM	*/*	*/*	*/*
OM	*0.00*	*/*	*/*

Table 14: *NSE values of the calibrated agent-based model on an independent dataset. Green values indicate superior performance than the uncalibrated model. Brown values indicate worse performance than the uncalibrated model.*

NASH_VALIDATION DATASET	1982	1990	1999	2007	2013
C	**1.00**	**0.81**	**0.74**	**0.63**	**0.42**
Y	**1.00**	**0.88**	**0.80**	**0.65**	**0.53**
M	**1.00**	**0.92**	**0.84**	**0.63**	**0.45**
O	**1.00**	**0.88**	**0.88**	**0.60**	**0.60**
T	**1.00**	**0.60**	**0.57**	**0.47**	***-0.34***

A spatial analysis of the model performance trough maps indicates the model indeed performs better compared to a null model, random model or model making use of the observed mean, but also shows the model is far from perfect. The direction of the net population evolution between 1982 and 2013 is correct predicted by the model in 66% of the villages (figure 59; table 15). The KIA indicates the model significantly outperforms a chance model, but this doesn't mean the model is perfect or very accurate because in 34% of the villages the population between 1982 and 2013 is modelled in the wrong direction. The performance of the model advances when the simulation period is shortened. The accuracy and KIA decrease with increasing simulation period (table 17). Figure 60 provides more information about over and underestimations of the population prediction. From 1982 to 1999, the predicted population is in >80% of the villages within 30% of the real population, but when the simulation is extended to 2013, this decreases to <60%. As mentioned, in some villages the population evolves in the wrong direction. In another part of the villages with a high deviation of the predicted population from the real population, the direction of the evolution is correct, but the increase in population is overestimated in some villages with an increasing population and the same goes for villages with a decreasing population.

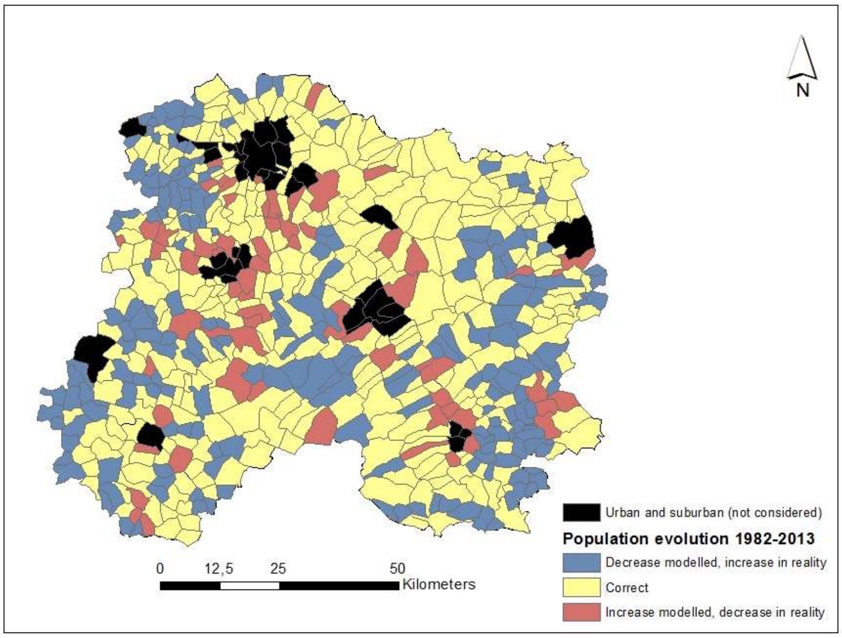


Figure 59: Modelled direction of the population evolution compared to observed direction of the population evolution. Own processing.

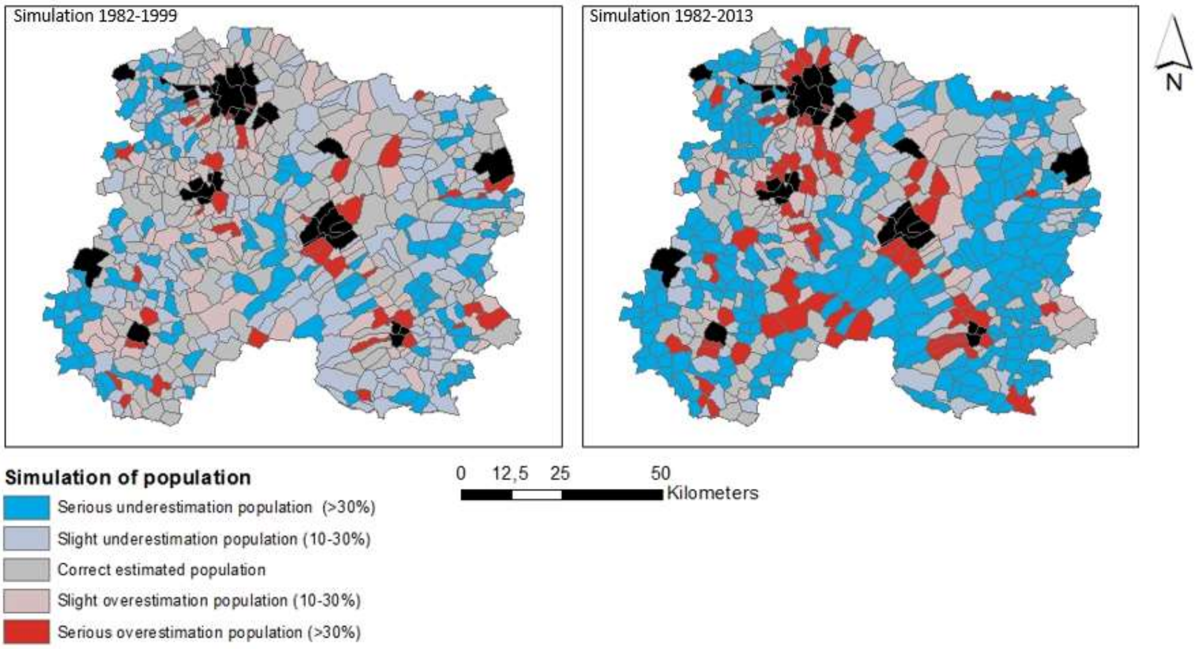


Figure 60: Under and over estimation of the population in the individual rural villages for (left) a simulation of the population in 1999 (simulated from 1982) and (right) a simulation of the population in 2013 (simulated from 1982). Own processing.

Table 15: *Confusion matrix for the simulation period 1982-2013. Indicates in how many villages of the Marne department an increase or decrease of the population is correct predicted.*

		Predicted	
		Decrease	Increase
Observed	Decrease	138	58
	Increase	140	245

Table 16: *Confusion matrix for the simulation period 1999-2013. Indicates in how many villages of the Marne department an increase or decrease of the population is correct predicted.*

		Predicted	
		Decrease	Increase
Observed	Decrease	148	53
	Increase	133	245

Table 17: *Accuracy and Kappa index of Agreement for different simulation periods for the prediction of the direction of the population evolution.*

Simulation Period	Accuracy	Kappa Index of Agreement
1982-2013	0.66	0.32
1999-2013	0.68	0.36

4.2.2. Model simulations and population projections for the future

The population graphs of some individual villages are shown (figure 61, figure 62). This are all villages visited during field work. The reason for the choice of this villages, is that they are more easy to be interpreted, because more knowledge about the dynamics of those villages is present. The exact locations of the villages are indicated on figure 13. The graphs shown are absolutely not the villages were the model performs the best, it are the most interesting villages to discuss in order to understand some phenomena or processes.

In **Congy** (figure 61A, figure 29), the population evolution is roughly correct predicted. The trend is fully correct and the absolute deviations deviate only at maximum 20 people. As discussed earlier, Congy is a rural village characterized with a high share of agricultural employment, explaining the continuous population decrease at a slow rate. When the model is initiated in 1982 and simulates the population until 2013, the prediction for this village is more accurate then when the model is initiated in 1999 and simulates the population until 2013. For the future, the population will stagnate or even slightly increase. There is not direct an obvious reason why the population trend would flip in the near future, however, during the conducted field work, the mayor of Congy was interviewed and he expected a stabilization or increase of the population in the future, but he couldn't argue why. A possible explanation of the modelled inverting evolution is the UL has as slower decrease compared to other villages in the region with a decreasing UL. This can lead to a flip of Congy from the villages with a UL below the median to the category of villages with a UL above the median. This implies less emigrations because the majority of emigrations happens in villages with an UL below the median.

In **Etoges** (figure 61B and figure 36) the direction of the population is wrongly predicted. The reason for the rise in population in Etoges is caused by a government intervention, the parceling of new pieces of land, leading to new sources of building land at low prices. This is an example of a village where it is logical to understand that a model isn't able of grasping this event. Abrupt events such as policy interventions cannot be predicted by this model. The consequence of the land parceling was an increasing population, initiating the rise of the SL of the village and also to a lesser extent stimulating local employment, slightly increasing the EP. This results in a higher UTS in 2013, and therefore the model with the data from 2013 as input data predicts a continuation of the soaring population.

In **Pocancy** (figure 61C and figure 30) the direction of the evolution of the population (1982-2013) is correctly predicted but the decrease in population happened in reality at a slower rate. The shorter simulation from 1999 to 2013 is more correct compared to the 1982-2013 simulation. Population is expected to decrease at a slow rate until 2025, after which the population decrease will accelerate according to the ABM.

In **Oyes** (figure 61D) population decreased slowly from 1982 until around 2010 after which the population decrease accelerated. The population in the village is expected to decrease further at this accelerated rate and in 2040 the village will become nearly completely abandoned. However, the real decrease of population between 1982 and 2013 is somewhat lower than predicted by the model.

In **Binson-et-Orquigny** the population remained almost stable between 1982 and 2013 (figure 62E). Binson-et-Orquigny is an isolated remote village without nearly any means of employment. The only service in the village left is a school. This school is a possible explanation the population is about stable and does not further decrease. The population is expected to remain stable for the next decades as well. In the case the school will close in the future, the population trend can possibly flip from stable to decreasing. Because in most villages is observed that all services close if the population decreases below a threshold of 300 to 250 people, there is a real chance the school will close in the future. In that case, the population is expected to decrease: figure 62F shows a simulation of the population in Binson-et-Orquigny with a hypothetical closure of the school in the year 2020.

In **Pogny** (figure 62G and also figure 33) , the population started to rise after the opening of a new commercial center around 2000. As is understandable, the model was not able to predict the opening of this commercial center.

In **Ecury-Sur-Coole** (figure 62H), a rural villages not that far from the regional city Châlons-en-Champagne (10-15km), the increase in population is seriously overestimated by the model. The obvious cause for this is its location not far from Châlons-en-Champagne (capital of the Marne department), implying the EP of the village is high or very high.

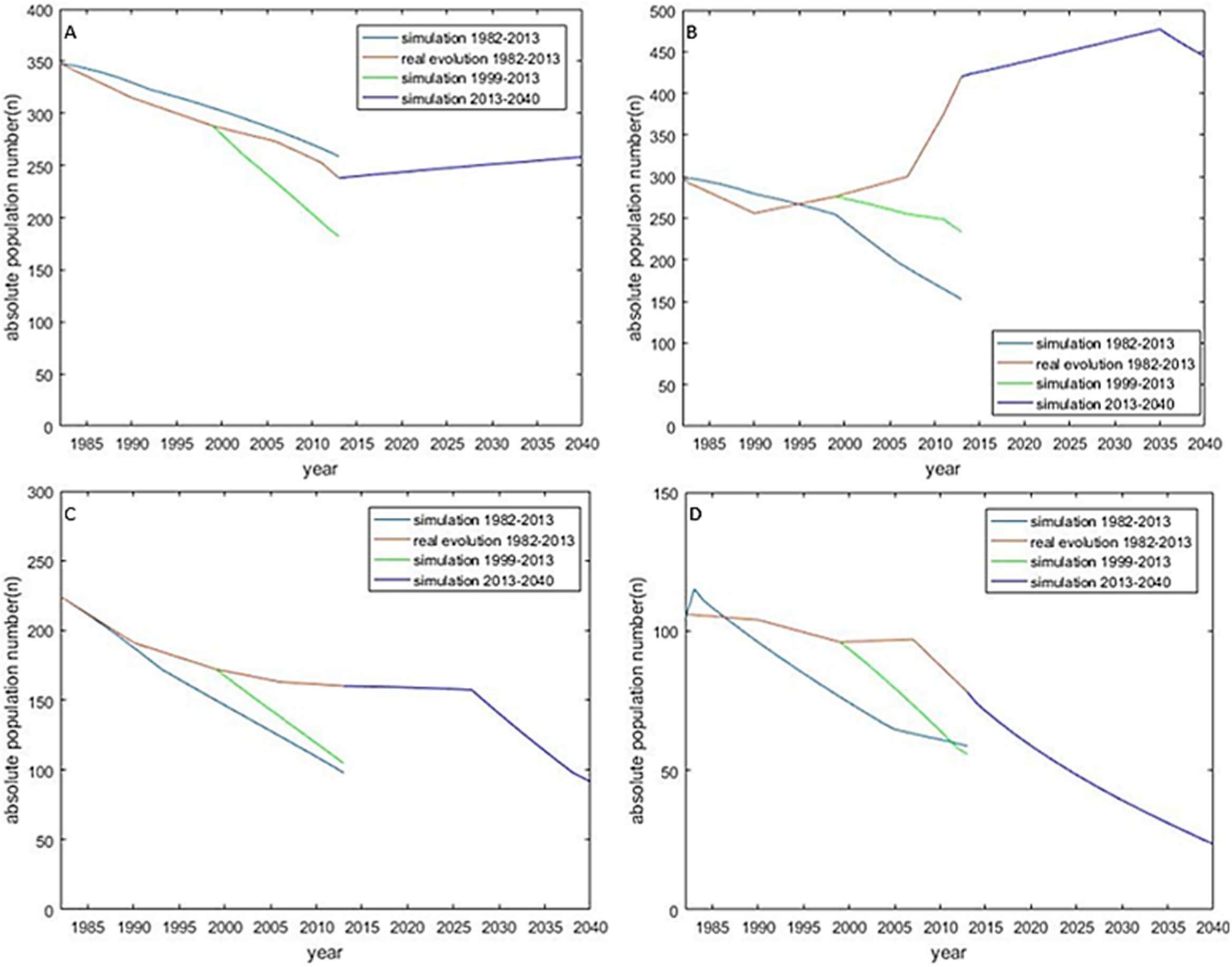


Figure 61: Observed and simulated population trends in a number of rural villages visited during field work. A= Congy, B= Etoges, C= Pocancy, D=Oyes. Location of the villages is indicated in figure 13 (chapter 2). Own processing.

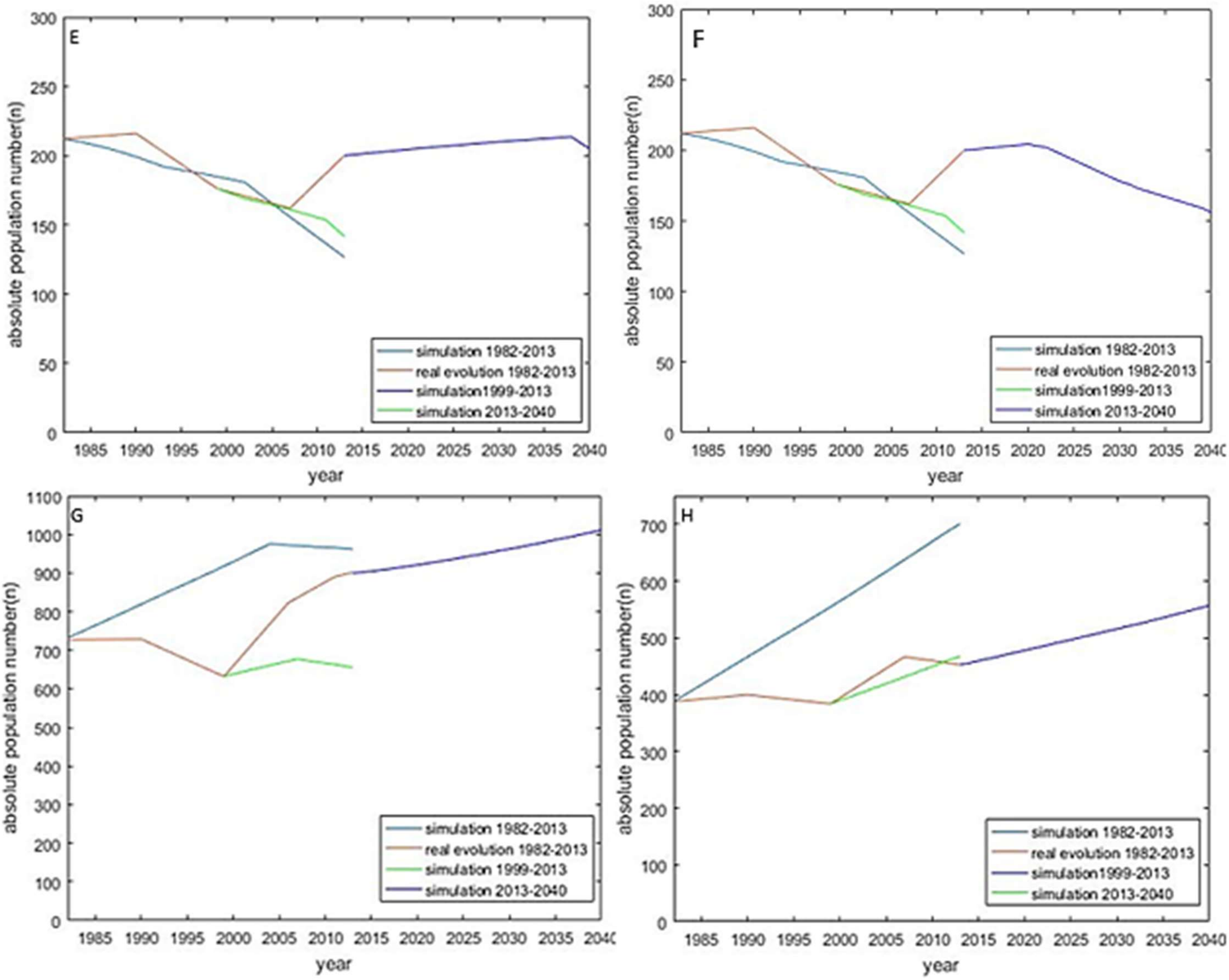


Figure 62: Observed and simulated population trends in a number of rural villages visited during field work. E= Binson-et-Orquigny, F= Binson-et-Orquigny with hypothetical school closure in 2020, G= Pogny, H= Ecury-sur-Coole. Location of the villages is indicated in figure 13 (chapter 2). Own processing.

The rural population in the Marne department will first increase very slightly until 2020, after which the rural population starts to decrease a little (figure 63). But this almost constant rural population masks high dynamics within the Marne department. Looking at the villages individually, the population tends to decreases in many of these villages and increases in other villages. The predictions for 2025 show that especially in the eastern villages of the Marne department huge population losses are to be expected, whilst the population will increases in the northern and central villages of the Department. In the southern region there are both villages with an increasing and a decreasing population (figure 63). Looking further into the future, to 2040, the majority of villages have decreasing population. Also in the regions with the highest increases in population like in the northern region around Reims, the population trend inverts and population decreases. There remain certain villages where the population rises spectacular. The location of those villages is somewhat spread across the department without an apparent spatial pattern.

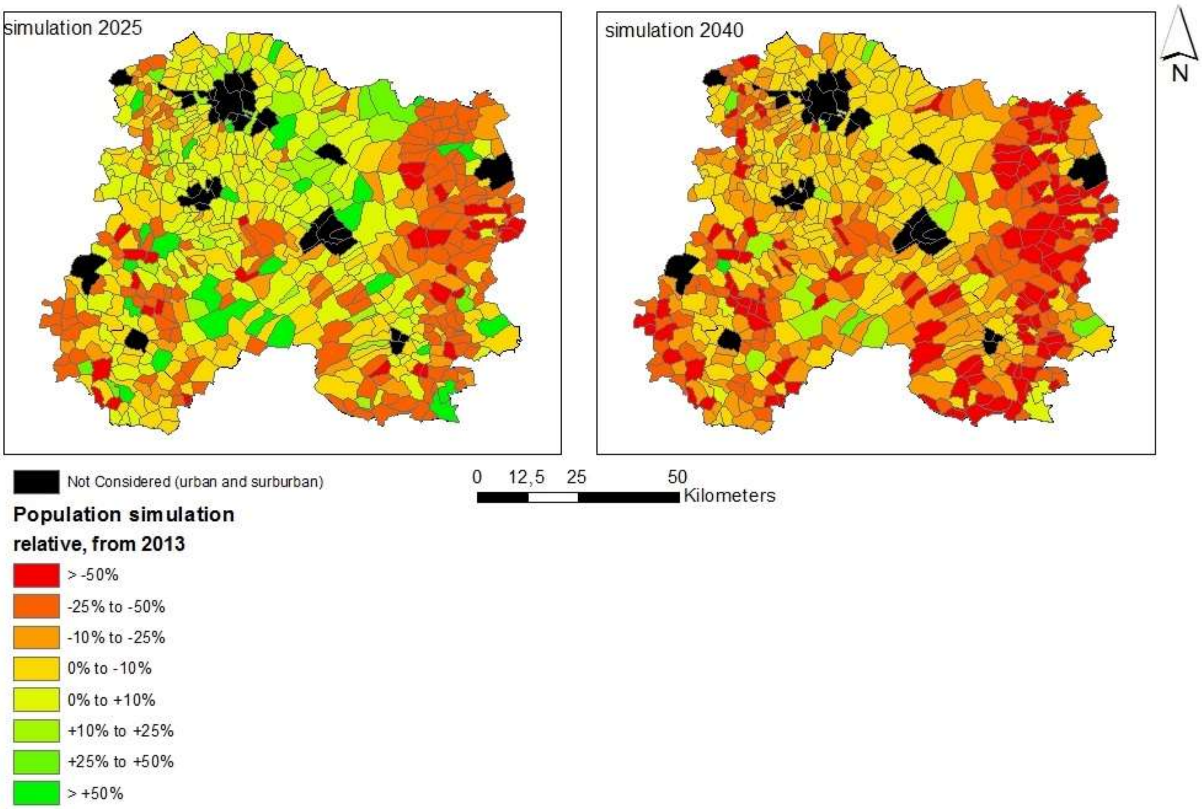


Figure 63: Simulated population evolutions from 2013 to 2025 (left) and from 2013 to 2040 (right). Own processing.

CHAPTER 5: DISCUSSION

5.1. Analysis of rural population dynamics

From a global perspective, the evolution of the fertility rate evolved downwards in the so-called developed countries first. The same phenomenon is also observed in the developing countries, but with a time lag. The fertility in France is analogue, first in the more developed regions a decrease is seen and later on also in the other regions, which were often characterized by a high percentage of rural population, few cities and a remote location. Despite the downward trend in the European fertility rates, France has one of the highest fertilities in Europe. For the future, a further decrease of the fertility rate in Europe is expected (van Nimwegen & van der Erf, 2010). This will lead to a net population decline in the largest parts of Europe (and France) in a few decades. Meanwhile the international migrations to Europe are expected to increase in the next decades, partially compensating a decreased natural growth of the population. However people who migrate from another country (inside or outside Europe) to an European country chose a place of residence in the cities in nearly all cases. This means that this probably not will affect the rural population dynamics. As a consequence, population decrease in rural areas in Europe that already experience depopulation can be supposed to accelerate. The Marne department is expected to follow the same trend as Europe, but the absents of urban area will probably lead to a more noticeable net decrease of population. In the past, the Marne department had a high fertility rate compared to France, but nowadays it is similar. It is worth noting that the fertility followed the same trend as in France but with a time lag. There was a fertility drop in the Marne in the '90s, but for France as a whole and the majority of French departments this decrease has already occurred before 1980. Within the Marne differences can be seen between the more urban and more rural areas. A higher fertility in the rural areas was found compared to the urban areas. The death rate was the same for the rural and the urban. Other existing researches came to the same findings, with systematically higher fertility rates in rural areas (Kulu, 2013). Various explanations can be given for higher fertility in rural areas: a higher proportion of poorly educated people in rural areas, the preference to raise children in detached houses (far more present in the rural areas), and the higher cost for raising children in the cities (Kulu, 2013).

Related to international migrations, due to the lack of immigrations to the Marne, the department will likely experience an accelerated population decrease. At present, there is already a net emigration from the Marne department. In particular, young people (19-34yr) often leave the Marne department. But the net emigrations are compensated by a natural increase of the population. If the fertility in the Marne department follows the same trend as the European average, natural growth of the population decreases and will contribute less to a population increase.

Within the Marne department, the number of rural-urban migrations outweighs the number of urban-rural migrations. Characteristics of the population also play a pivotal role in the occurrence of migrations within the Marne: residential mobility within the department is far higher for younger people (19-34yr) compared to older people, of which the majority is non migrant. Certain studies have concluded that residential mobility depends on received education (Détang-Dessendre et al., 2008; Machin et al., 2012). Those characteristics of the population can also to a certain extent influence the characteristics of a village, region or country. The average residential mobility is higher in countries where people are on average better educated (figure 64).

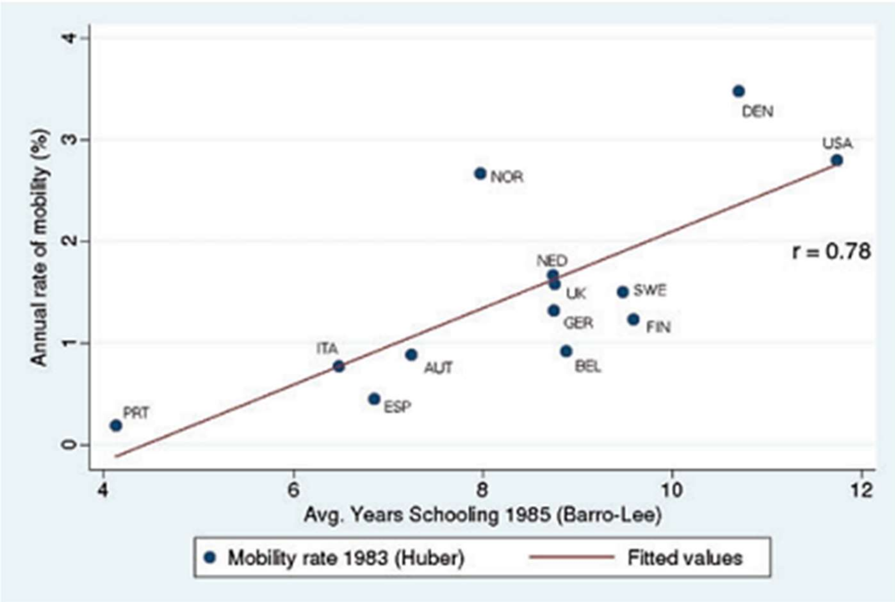


Figure 64: Residential mobility in function of years of education (Machin et al., 2012)

Beside residential mobility, the typology of migrations can also be explained by education level and age. Highly educated people migrate more often rural-urban while for the moderately educated people rural-rural migrations are the most important. Poorly educated people almost never migrate towards urban areas: rural-rural and urban-rural migrations occur at an equal amount in this group. Regression analysis demonstrated age and population combined do not better predict migration typologies compared to education alone. However, in some cases age rather than education does play a role, such as young people leaving their parental house or couple-forming and the birth of children (Détang-Dessendre et al., 2008). But this research revealed that education rather than age can more often explain migration types. This is perhaps one of the groundbreaking revealing of this research. Existing literature investigates the relationship between urban and suburban residents and the migration types. To the best of the researchers knowledge, however, no single research investigates the relationship between the typology of migrations from rural areas and received education. This illustrates that it is often impossible to derive cause-consequence interpretations from regression analysis.

Both existing studies in literature and the agent-based modelling conducted in this research indicate that the population will decrease in the majority of the rural villages of the Marne department in the next decades (INSEE, 2010). In the majority of the rural villages located in eastern and southern regions of the department, serious depopulation has already occurred, and the future depopulation is expected to be the most severe in the villages in those regions. On a smaller scale, there are also differences within each region: in areas with a rising population, some villages will decline and vice versa some villages in depopulating areas will survive or even grow. A lot of factors control this phenomenon to explain those differences between individual villages. In this research, significant correlations were found between the evolution of the population in rural areas and the EP, the SL (and distance to the supermarket DTS), the HP, the elevation and the distance to a city. Elevation was the only factor related to the physical environment that showed a significant correlation with the population evolution in rural areas. Other research found also the distance to the sea as an important factor (Baccaïni & Dutreuilh, 2007; Di Figlia, 2016). In this research, however, is this distance to the sea too large to have an effect.

Elevation as a variable related to the physical environment can play a role in migratory decision in the Marne department, and the three most important factors controlling rural population dynamics in the Marne department are more socio-economic: the EP , the SL and the HP. In certain studies, employment is cited as an important factor, as are housing prices (So et al., 2001). The services in a village are often seen as a consequence of depopulation, but there seem to exist no studies taking into account this as causal variable. What from this research appears, is that the SL is mostly not the initial cause of depopulation. In a large majority of the villages, the SL initiated decreasing after the population fall. Some rare exceptions were found where the population started to fall after the SL decreased (e.g. Châtillon-sur-Marne). However the SL is mostly not the initial cause of population decrease, it certainly accelerates depopulation when the process once started. This can be called a “positive feedback[8]”.

However the employment is often mentioned in other existing studies, the implementation of the employment in the model is quite innovative compared to previous researches. In the majority of existing studies, either distance to a city or only the number of local jobs in the village itself are under consideration, while in this research an EP is calculated taking into

[8] A positive feedback is a process that enforces an ongoing process.

account both the employment in the village itself and the employment in the surroundings of the village. Also a distinction is made between employment in the agricultural sector, industrial sector and service sector. The result from this analysis is that during the last decades the EP and the population decreased mainly in villages were agricultural and industrial employment were dominant. Rural villages with a high percentage of rural employment were mostly characterized with a continuous population loss at a more or less constant rate while villages characterized with a high industrial employment often experienced an abrupt population fall some years after the disappearance of industrial jobs in the surrounding areas when a factory closed. Villages were the EP increased the last decades had often already a high percentage of service employment at the commencement of the analysis in 1982. What is discussed is the case in the majority of the rural villages, there are also some exceptions: in a small fraction of the villages industrial employment increased the last decades (figure 44) and in a few villages the EP increased even if there was a high percentage of industrial or agricultural employment (compare figure 43 and figure 44). The calculation of the EP can be further improved by taking not only into account the distance to the job location but also the available transport and quality of the transport infrastructure to reach the place of employment.

The HP are also an important consequence of population evolutions: while those prices are higher in the villages with increasing population, they can decelerate depopulation, because the HP fall in depopulated villages.

The relative importance of those factors (EP, SL, HP) depend on the age of the individual. Therefore, the characteristics of population and villages are also interlinked: in villages where mainly young people leave, the remaining population is elder. This implies also less births in this villages, which leads towards a lower natural growth of the population in those villages.

Based on those considerations, the rural population of the Marne department is expected to decrease in the next decades because the natural growth of the population will slow down. At present the number of emigrations outweighs the number of immigrations and there are no indications this will change in the near future. Thus, in general the rural population will decrease but there will be a spatial pattern at different levels: some regions where the SL, EP are low such as in the eastern and southern part of the Marne department are expected to experience more rural depopulation than the region around Reims. Sometimes rural depopulation can cause that village become nearly or completely abandoned, this are called ‘ghost villages’. The most important findings about rural abandonment are discussed in the next section.

5.2. Rural Abandonment

Rural abandonment clearly occurs in the Marne department and causes problems. Most of the problems listed in the literature review of rural abandonment (chapter 1) were observed on terrain during field work (chapter 3). Closing of services such as shops and recreational activities, the shutdown of public transport, lack of employment or long commuting distance to the employment place and social exclusion are the most important observed problems. Closure of services seems to happen after the population falls below a certain threshold located around 250 to 300 people. The shutdown of public transport depends often not on the population in one village but the sum of the population of some villages around a transportation trajectory. Not all the consequences of rural abandonment are negative. In certain depopulating areas, sometimes reforestation of grassland and to a lesser extent arable land is observed. This is good for ecosystem functioning and enhances biodiversity (Derh et al., 2016).

Which rural villages will experience rural abandonment or in which rural villages already existing rural abandonment will worsen depends on the controlling factors in the village and the area, but in general, the problem is expected to worsen in the Marne department. This also appears from the agent-based simulation of the population evolution in the next decades for the rural villages. The worsening of rural abandonment probably will not be limited to the Marne department, but will probably also happen in other French regions and in other regions and countries in Europe. Arguments for this is that similar process of rural abandonment are also occurring in other European countries and there are no indications regarding whether there would be an abrupt strong differentiation in the future, however expanded research is needed to explore these trends.

Policy initiatives to prevent or slow down rural abandonment exist, as at European level, national level as local level. The European policy dominantly focusses on agriculture with direct payments to farmers to get started or subsidies to produce certain cereals. The European policy regarding rural areas has various critics: for example, some say there is too much focus on agriculture and too little focus on other aspects of rural development. It can be questioned whether those investments in agriculture are sustainable and should be continued or not. Farmers are ageing, there is far more food produced in Europe than needed for the European citizens, the European agricultural policy causes economic problems in some developing countries because certain highly subsidized agricultural outputs are dumped on their market at very low prices causing troubles for local farmers there and the policy causes high biodiversity losses. It is also naïve to expect a stabilization or an increase in agricultural employment in Europe in the next decades. The world is continuously changing and some trends are irreversible. The reality is that trough technological developments less and less manpower is needed to sustain satisfactory levels of food production in Europe.

It can be an idea to allow a small fraction of the rural villages in depopulating areas to collapse and then convert those areas to a natural habitat, such as forest. In that scenario, the last remaining population in those villages could be given subsidies to move elsewhere or to offer them cheap places in flats, service flats or other buildings. Those flats or service flats could be built in certain rural villages with an increasing population. This implies that the remaining people, mostly elderly people, become inhabitants of a rural village (they do not have to move to a city if they do not want that) with many services. Also it would be possible to convert some rural villages to ecovillages or tourist destinations (discussed below). Another strategy can be to invest equally in all rural villages but that would have a small chance to succeed. It is not very plausible the general process of rural depopulation can be stopped or reverted, the phenomenon is observed in more and more places in the world and is becoming worse and worse in the places where it occurs.

On the national level, there are also policies related to rural development, as in France as in other European countries, ranging from tax cuts to start-up companies to programs for better education. Other policies seen in certain European countries related to rural abandonment are the shutdown of public transport. In a lot of regions of the Marne department, all the public transport is shut down, causing problems for people who can't afford a car or for elderly people who are not able to drive themselves anymore. This is an illustration of a policy intervention enhancing rural abandonment and accelerating rural depopulation. On the local level, also measures to prevent rural abandonment are taken, such as parceling of new pieces of land so that such land will become inexpensive areas on which to build houses.

Examples of successful revitalization of certain rural villages through policy interventions can be found in Italy. These revitalization programs succeeded to either convert an abandoned rural village in a popular tourist destination or in certain other villages to renew the human activity by converting the villages to 'ecovillages' (Di Figlia, 2016). By doing this, these programs have aimed at making those villages an attractive place to live. All of those options have advantages and disadvantages: converting the rural village to a tourist destination is good for the development of the local economy and creates jobs for the local people, but it is less favorable considering environmental protection and the preservation of heritage (Di Figlia, 2016). Ecovillages on the other hand are favorable from an environmental perspective, but fewer jobs are created than when the village is converted into a tourist destination (Di Figlia, 2016).

> *"Common to ecovillages are shared facilities for: community meals, meetings, and other activities, laundry, and recreational spaces like playgrounds and swimming areas"* (Kasper, 2008).

5.3. The added value of an agent-based model

The approach of using an agent-based model was quite innovate because no single other study is found where the population evolution was simulated with this technique and the simulation was validated. There seems to exist only one study about agent-based modelling of human population (Zhang & Jager, 2011), but in that study nothing about model validation or calibration is mentioned. So from that research it was not clear whether agent-based modelling can be suitable for population simulations.

For the considered period 1982-2013, the performance of the agent-based model is clearly superior compared to the performance of a chance model, a null model and a model using the observed mean of the simulation period to simulate the population evolution. Also was found the model also outperforms a stepwise regression model (developed in this research) where the population evolution was simulated using regression equations. This regression model took into account EP and SL, the same factors being the most important in the agent-based model. The superior performance of the ABM can be explained because a lot of factors included in the ABM couldn't be included in the STR. In the STR the population is simulated only using explanatory variables. Characteristics of the population such as the preference of different migration types depending on the agents age and other things like the annual migration rate, importance of services according to their own judgement, the birth rate, the death rate and the type of employment weren't included in the STR. This is probably one of the plausible explanations for the superior performance of the ABM. Another explanation is that more complex relationships (e.g. not only linear relationships) between variables can be incorporated in the ABM.

However, the performance of the ABM is far from perfect with in about 30% of the villages the population evolution simulated in the wrong direction. For certain villages, where the direction of the population evolution is correctly predicted, the population is overestimated in the case of an increasing population and underestimated for villages with a decreasing population. There is probably an explanation for this: the HP are not included in the model because no reliable data about the HP were available. This is because, as discussed, the HP decelerate depopulation in villages with decreasing population and decelerates population growth in the villages with increasing population. Therefore, over- and under-estimations would be reduced if it was possible to include the HP in the model.

In general, the model could also further improved by also including the education level of the agents and the elevation of the center of the village. The education level is very important in explaining which people will make which migratory decisions. But, as with the HP, a lack of reliable data is the cause that the factor is not included in the model. The elevation is less important but it is possible that it improves the model performance slightly. Also data extracted from existing policy plans related to the future development of the regions can possibly improve performance.

CHAPTER 6: CONCLUSION AND SCOPE FOR FURTHER RESEARCH

Comprehensive analyses were accomplished to obtain insights in the process of rural population evolutions in the Marne department. The core analyzes was focused on spatial and statistical analysis of large amounts of data obtained from various sources such as field work and existing databases. Further, it is demonstrated agent-based modelling can be effective in simulating population evolutions. This conclusion provides the answers on the research questions stated at the beginning of the research.

Similar patterns in demographic evolutions can be observed at different spatial scales:

On the level of the individual villages. There is a clear spatial pattern in the population evolution in the rural villages of the Marne department. Villages with a decreasing and increasing population are to a certain extent clustered together. There is a clear axis of villages with increasing population. This axis is located between Reims and Châlons-en-Champagne, constituting the valley of the Marne river, located at the lowest elevation of the department. The main road and railways are situated along this axis as well. The dominant land use in this region is arable land. Villages with a decreasing population are especially seen in eastern and southern regions of the Marne department. Those villages are often more remote and characterized by an already low population density. The land use in these regions is a combination of forest and arable land.

On the level of the departments. The drivers of the population evolutions are to a large extent the same as for the individual villages. However, for the departments also the distance to the sea has an influence. All coastal department have an increasing population while all departments with a decreasing population are located inland.

When people are dissatisfied with their place of residence, some of them will migrate. The most important **push factors** related to the village characteristics are low EP for young- and middle-aged people and a low SL for elderly people.

Once people have decided to migrate, they need to choose a new village of residence. The most important **pull factors** for this are a high EP, a high SL, and low HP. For young- and middle-aged people the EP is the most important. For retired people, the SL is the most important consideration. Also the income level can play a role, especially for poorer people: some people are willing to migrate to a place where they expect a higher satisfaction (villages with higher UL), but are not in the possibility to finance this migration, because there are often higher HP at places where individuals experience a higher general satisfaction related to their place of residence. For those people, the pull force of the low HP is higher than the pull force of the high EP and SL. The HP often reflect the state of the rural villages with a high HP in the rural villages in good condition and a low HP in the rural villages that are deteriorating. Related to the EP, not only the amount of employment opportunities is important. Also the type of employment matters: villages with a high percentage of employment in the service sector have a higher probability to thrive than remote villages where employment in industry or agriculture is dominant.

Less important, but not negligible, is the role of the physical factor elevation in the population evolution. Despite the limited range of the elevation within the Marne department, there is more depopulation observed in rural villages of which the center is situated at a higher altitude.

Regarding the characteristic of the population, well-educated people and young people are generally more eager to migrate. The least eager to migrate are older people and especially people who only graduated from primary school. It can be stated that the push and pull factors related to the village typology should be stronger for elderly and lesser educated people to migrate than for younger and better educated people. There is a cause-effect relationship between residential mobility and the education received. Not only the quantitative aspect of migrations is explained by education and age, but also the type of migrations a person decides to make. Especially well educated people with an university degree often migrate from the rural to the urban. People without any kind of degree besides a primary school degree often migrate from urban to rural places (when they initially live in the urban) or from the rural to the rural (when they initially lived in the rural). Average educated people more often opt for rural-rural migrations. Also age explains to a certain extent the type of migrations people consider (mainly more rural-urban for younger people), however this is less important than education.

The phenomenon of **Rural abandonment** occurs in places where the UL is the lowest, and rural depopulation is a self-accelerating process: if people leave the village, over time, services start to close when the population falls below a certain threshold. In those villages, the rural depopulation process can evolve that far that certain villages can become almost or totally abandoned. This has many undesired consequences for the people that stay such as social exclusion, increased remoteness, in combination with the closing of services in the villages such as the public transport, shops and recreational activities. For the active population, also a lack of employment and/or long commuting to the place of employment are often problematic in these villages.

Rural abandonment is absolutely not limited to the Marne department, but also takes place in other regions of France, Europe and the World. In many areas the problem is expected to worsen over the next decades. However, there are also some examples that demonstrate it is possible to deal successful with the problem. It is doubtful whether investing in all rural villages can be a success story. It is recommended to not preclude other scenarios in which the spatial structure changes. Examples are converting a fraction of the deteriorating villages to tourist destinations, ecovillages or nature and move a fraction of the population in strong deteriorating villages to rural villages with an increasing population. It might be useful to investigate the feasibility and the support for these scenarios.

The next goal of the research was finding out whether an **agent-based model** can be suitable for simulating population evolutions. The answer to that question is yes, and it is significantly more accurate than a null model, chance model, a model using the observed mean during the simulation period and a stepwise regression model. The agent-based population simulations can also be useful to predict the evolvement of the problem of rural abandonment in the future. The agent-based model developed in this study predicts the rural abandonment in the Marne department will worsen. It predicts that by 2040 a large majority of the rural villages will experience rural depopulation. In about one third of the rural villages is even simulated the population will more than decimate between 2013 and 2040.

However, there are many possibilities to expand the agent-based model which can lead to an even more accurate version of the model. Possible **scopes for further research are**:

(I) Add additional data to the model**:** The available research tools were limited in this research. The research is conducted by a single researcher. As a larger research team would conduct field work for several weeks, of course far more data can be gathered. During fieldwork all the interviewees were asked about their education level and that of their relatives, but there were no public databases on education levels available and the dataset of the fieldwork was to limited to extrapolate. Therefore the factor could not be included in the model. Also inclusion of reliable HP and the elevation of the center of the villages possibly can improve model performance.

(II) Expand the spatial extent of the model: In literature many studies about rural abandonment for various areas of the world exist. The approach in this study is different than the approach used in existing studies. Therefore, this can be an introductory for further research. The model can be expand to other similar rural regions in Europe where the problem of rural abandonment is present (e.g. other French departments such as the Ardennes, Haute-Marne and Centre, other countries like Bulgaria, Estonia etc.), with only relative limited fieldwork time. The model could also be tested in regions where rural depopulation is slowed down or reverted in some villages like some Italian regions where certain villages are converted to tourist destinations or ecovillages. This can be useful to see how the model can be adapted to project the influence of some of these measures.

(III) Include big data in the analysis: In the domain of population modelling, agent-based modelling can be considered as a relatively new and innovative approach. Combining agent-based modelling with even newer approaches can result in improved and more accurate population simulations. For example big data can be included such as the online surfing behavior in villages. If in certain villages, terms as moving away, migrate, are often searched, it can indicate there are (many) people considering moving away from the village. Another possible example is the spatial pattern and movement of mobile signals during day which can be an indicator of the commuting distance people of the village have to travel to their employment place. If factors such as the aforementioned are combined with agent-based modelling, it is realistic to expect very accurate population simulations.

(IV) Include remote sensing data in the analysis**:** Information about the fraction of green surface and built-up surface in the villages can possibly have an added value. It can be also useful to know the number of houses in each village and compare it to the population number of the villages. This can provide information about the number of vacant houses in the village. Information regarding the space between houses can improve delineation of rural and suburban villages.

Rural abandonment is a real problem and is expected to worsen. This has many undesired consequences for the remaining population in the deteriorated villages. If no or insufficient action is taken, the abandonment and the negative consequences for the people will probably worsen. But there is also a scenario possible whereby a number of problems related to rural abandonment can reduce. Therefore, two conditions should be fulfilled. First, enough tools and resources are needed in combination with enough researchers who are willing to keep up to the problem. A second prerequisite is the political willingness to look at the problem from a perspective of how the world is today and not from how the world was yesterday. This implies less focus on agriculture and more focus on rural development and spatial planning. If those conditions are met, there can be hope that many problems related to rural abandonment and the resulting problems such as a decreased quality of life for the remaining population in depopulating villages, can disappear.

REFERENCES

Journal Articles & Books

Antrop, M. (2004). Landscape change and the urbanization process in Europe. *Landscape and Urban Planning*, *67*(1–4), 9–26.

Azarnert, L. V. (2014). Integrated Public Education, Fertility and Human Capital. *Education Economics*, *22*(1–2), 166–180.

Baccaïni, B., & Dutreuilh, C. (2007). Inter-Regional Migration Flows in France over the Last Fifty Years. *Population (English Edition, 2002-)*, *62*(1), 139–155.

Baccaïni, B., & Levy, D. (2009). Recensement de la population de 2006. *Première*.

Balar, D. J. (2012). Report Information from ProQuest, (January), A.2.

Boysen, O., Jensen, H. G., & Matthews, A. (2014). Impact of EU agricultural policy on developing countries: A Uganda case study. *The Journal of International Trade and Economic Development*, *8199*(July), 1–34.

Brandmülle, T., Lupu, I., Önnerfors, Å., Corselli-Nordblad, L., Coyette, C., Johansson, A., … Wolff, P. (2016). *Urban Europe*.

Castillo-Montoya, M. (2016). Preparing for Interview Research : The Interview Protocol Refinement Framework Preparing for Interview Research : The Interview Protocol Refinement. *The Qualitative Report*, *21*(5), 811–831.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250.

Comber, A. (2008). Land use or land cover? *Journal of Land Use Science*, *3*(4), 199.

Dahiru, T. (2008). P-Value, a true test of statistical significance? A cautionary note. *Annals of Ibadan Postgraduate Medicine*, *6*(1), 21.

Dammers, E., & Keiner, M. (2006). Rural Development In Europe. *disP - The Planning Review*, *42*(166), 5–15.

De Haas, H. (2011). Mediterranean migration futures: Patterns, drivers and scenarios. *Global Environmental Change*, *21*(SUPPL. 1), S59–S69.

De la Croix, D., & Dottori, D. (2008). Easter Island's collapse: A tale of a population race. *Journal of Economic Growth*, *13*(1), 27–55.

Derh, M. A., Murphy, H., Monteith, G., & Men??ndez, R. (2016). Measuring the success of reforestation for restoring biodiversity and ecosystem functioning. *Journal of Applied Ecology*, *53*(6), 1714–1724.

Détang-Dessendre, C., Goffette-Nagot, F., & Piguet, V. (2008). Life cycle and migration to urban and rural areas: Estimation of a mixed logit model on French data. *Journal of Regional Science*, *48*(4), 789–824.

Di Figlia, L. (2016). Turnaround: abandoned villages, from discarded elements of modern Italian society to possible resources. *International Planning Studies*, *21*(3), 278–297.

Eea. (1994). CORINE land cover - contents. *Methodology*, 1–163.

European Commission. (2008). *Poverty and social exclusion in rural areas*.

European Commission. (2013). Overview of CAP Reform 2014-2020. *Agricultural Policy Perspectives Brief N.° 5*, (December 2013), 10.

European Commission. (2015). DIRECT PAYMENTS FOR FARMERS 2015-2020.

Faraway, J. J. (2002). Practical Regression and Anova using R. *Reproduction*, *21*(July), 212.

Ferreira, J. C., & Patino, C. M. (2015). What does the p value really mean ?, *41*(5), 3713.

Forsyth, a. (2012). Defining Suburbs. *Journal of Planning Literature*, *27*(3), 270–281.

Gray, C., & Mueller, V. (2012). Drought and Population Mobility in Rural Ethiopia. *World Development*, *40*(1), 134–145.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: a guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350.

Huang, R., & Cui, H. (2015). Consistency of chi-squared test with varying number of classes. *Journal of Systems Science and Complexity*, *28*(2), 439–450.

Hunter, L. M. (2011). NIH Public Access, *21*(5), 16–21.

Hyndman, R. J., & Koehler, A. B. (2005). and Business Statistics Another Look at Measures of Forecast Accuracy Another look at measures of forecast accuracy. *International Journal of Forecasting*, *22*(November), 679–688.

INSEE. (2010). La population des régions en 2040, 4.

Kasper, D. V. S. (2008). Redefining community in the ecovillage. *Human Ecology Review*, *15*(1), 12–24.

Kesternich, I., Siflinger, B., Smith, J. P., & Winter, J. K. (2014). The Effects of World War II on Economic and Health Outcomes across Europe. *Review of Economics and Statistics*, *96*(1), 103–118.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, *15*(3), 259–263.

Kulu, H. (2013). Why Do Fertility Levels Vary between Urban and Rural Areas? *Regional Studies*, *47*(6), 895–912.

Lampard, E. (2016). The History of Cities in the Economically Advanced Areas Author (s): Economic Development and Cultural Change , Vol . 3 , No . 2 , The Role of Cities in Economic Development and Cultural Change , Part 2 (Jan ., 1955), pp . 81, *3*(2), 81–136.

Li, X. (2015). Rural Depopulation in China: A Comparative Perspective. *International and Multidisciplinary Journal of Social Sciences*, *4*(2), 149.

Liu, C., & Laske, E. (2010). The Empty Rural Area Diagonal in France (" Diagonale du vide "), *14*, 90–96.

Lu, X., Bengtsson, L., & Holme, P. (2012). Predictability of population displacement after the 2010 Haiti earthquake. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(29), 11576–81.

Macal, C., & North, M. (2010). Tutorial on agent-based modelling and simulation. *J Simul*, *4*, 151–162.

Machin, S., Salvanes, K. G., & Pelkonen, P. (2012). Education and mobility. *Journal of the European Economic Association*, *10*(2), 417–450.

Manakos, I., & Braun, M. (2014). *Land Use and Land Cover Mapping in Europe*. *Land Use and Land Cover Mapping in Europe* (Vol. 18).

Marshall, M. N. (1996). Sampling for qualitative research Sample size. *Family Practice*, *13*(6), 522–525.

Mayer, D. G., & Butler, D. G. (1993). Statistical validation. *Ecological Modelling*, *68*(1–2), 21–32.

McArthur, D. P., Thorsen, I., & Ubøe, J. (2014). Employment, transport infrastructure, and rural depopulation: A new spatial equilibrium model. *Environment and Planning A*, *46*(7), 1652–1665.

Mladenov, C., & Ilieva, M. (2012). The depopulation of the Bulgarian villages. *Bulletin of Geography. Socio-Economic Series*, *17*(17), 99–107.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Binger, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.

Mougel, N. (2010). World War I casualties. *Romania*, 1–16.

Nobles, J., Frankenberg, E., & Thomas, D. (2015). The Effects of Mortality on Fertility: Population Dynamics After a Natural Disaster. *Demography*, *52*(1), 15–38.

OECD. (2014). *New Rural Policy: LINKING UP FOR GROWTH*.

Parker, D. C., & Meretsky, V. (2004). Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. *Agriculture, Ecosystems and Environment*, *101*(2–3), 233–250.

Pfister, C., & Brázdil, R. (2006). Social vulnerability to climate in the “Little Ice Age”?: an example from Central Europe in the early 1770s. *Climate of the Past*, *2*(2), 123–155.

Poelmans, L., & Van Rompaey, A. (2010). Complexity and performance of urban expansion models. *Computers, Environment and Urban Systems*, *34*(1), 17–27.

Satterthwaite, D. (2007). The transition to a predominantly urban world and its underpinnings. *Human Settlements Discussion Paper Series: Urban Change-4*, (2006), 99p.

Schreinemachers, P., & Berger, T. (2011). An agent-based simulation model of human-environment interactions in agricultural systems. *Environmental Modelling and Software*, *26*(7), 845–859.

Sharma, M. J., & Yu, S. J. (2015). Stepwise regression data envelopment analysis for variable reduction. *Applied Mathematics and Computation*, *253*, 126–134.

Shcherbakov, M. V., Brebels, A., Shcherbakova, N. L., Tyukov, A. P., Janovsky, T. A., & Kamaev, V. A. evich. (2013). A survey of forecast error measures. *World Applied Sciences Journal*, *24*(24), 171–176.

So, K. S., Orazem, P. F., & Otto, D. M. (2001). The effects of housing prices, wages, and commuting time on joint residential and job location choices. *American Journal of Agricultural Economics*, *83*(4), 1036–1048.

Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing and Management*, *45*(4), 427–437.

Tan, P.-N., Steinbach, M., & Kumar, V. (2005). Chap 8 : Cluster Analysis: Basic Concepts and Algorithms. *Introduction to Data Mining*, Chapter 8.

van Nimwegen, N., & van der Erf, R. (2010). Europe at the Crossroads: Demographic Challenges and International Migration. *Journal of Ethnic and Migration Studies*, *36*(9), 1359.

Viera, A. J., & Garrett, J. M. (2005). Understanding Interobserver Agreement: The Kappa Statistic, (May), 360–363.

Wilson, S. G., & U.S. Census Bureau. (2009). Population dynamics of the Great Plains : 1950 to 2007. *Current Population Reports. Population Estimates and Projections*, (P25-1137), 19 p.

Yap, B. W., & Sim, C. H. (2011). Comparisons of various types of normality tests. *Journal of Statistical Computation and Simulation*, *81*(July 2015), 2141–2155.

Zhang, Q., & Jager, W. (2011). Agent based modeling of population dynamics in municipalities : Migration in the Derbyshire & Nottinghamshire cases in the UK . *Evolution*, *6*(November 2008).

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Annals of Translational Medicine*, *4*(7), 136.

Zhou, D. M., Xu, J. C., Radke, J., & Mu, L. (2004). A Spatial Cluster Method Supported by Gis for Urban-Suburban-Rural Classification. *Chinese Geographical Science*, *14*(4), 337–342.

Online references

BRGM, 2017.Bureau de Recherches Géologiques et Minières http://www.brgm.eu/. Last consulted 17/2/2017.

ESRI (2017). http://www.esri.com, last consulted 6/5/2017.

Eurostat (2016). Your key to European statistics, http://ec.europa.eu/eurostat/data/database. Last consulted 8/5/2017

Government France (2017). Information on classification in administrative zones. http://www.gouvernement.fr , last consulted 24/02/2017.

INSEE (2015). Récensement de la population. Data about population evolutions http://www.insee.fr , last consulted 4/3/2017.

INSEE (2016). Récensement de la population. Data about population evolutions http://www.insee.fr , last consulted 14/4/2017.

INSEE (2017). Récensement de la population. Data about population evolutions http://www.insee.fr , last consulted 29/3/2017.

Meilleurs Agents (2016). Prix immobilier. http://www.meilleursagents.com/prix-immobilier/ Last consulted 27/11/2016.

NASA (2017). National Aeronautics and Space Administration. Data about elevation. https://www2.jpl.nasa.gov/srtm/france.htm. Last consulted 16/2/2017

OECD (2016). Organization of Economic Co-operation and Development. Definition of suburban area. http://www. http://www.oecd.org/. Last consulted 5/5/2017.

Spatial Foresight (2016). Maps about population evolutions, migrations and natural growth of the population in Europe. http://www.spatialforesight.eu/ , last consulted 4/3/2017.

United Nations (2016) Population projections. Online available at: http://www.un.org/en/index.html, last consulted 8/5/2017.

Appendix A: Interview Questionnaire

Informations générales sur l'ensemble du ménage

Qui habite dans cette maison? (interviewé dans première rangée) Ques-ce-que l'âge , le niveau d'éducation (primaire/secondaire/l'ensiegnement supérieur/université) et l'occupation professionelle de gens qui habite dans cette maison?

	Nom	Age	Education	Profession (depuis quand)
1				
2.				
3.				
4.				
5.				
6.				
7.				
8.				

2. Information sur l'interviewé (parcours l'espace-temps de la naissance jusqu'à present)

Où avez vous grandi? Quand avez-vous déménagé dans ce village? En-quelle année? Où avez-vous vécu(/habite) avant de déménager à ce village? (Est-il situé loin d'ici? Est-ce un petit village ou une ville? ...). Pensez-vous déménager vers un autre endroit à l'avenir (proche ou lointain)? Quand? A ou ?

Grandi	
Déménagé dans ce village ? Annéé ?	
Vécu devant déménager	
Plans déménager dans la future?	
Pourqoi êtes-vous venu vivre ici?	

3. Emploi (tous les membres du ménage)

Êtes-vous employé? (If retired, go to 4) Quelle sorte de travail faites-vous? Où est votre lieu de travail situé? Quel type de transport utilisez-vous pour rejoindre votre lieu de travail? (Vélos, voiture, train, bus?) Ca prend combine de temps? Y at-il une évolution dans le temps de Voyage à votre lieu de travail (plus ou moins que dans le passé)

Employéé: oui non	Sorte de travail:

Lieu de travail:	Transport type:
Travel time and distance	Evolution:

Est-ce-qu'il ya plus ou moins d'emploi dans la region par rapport à il ya 20 ans?

4. Reconstruction de l'histoire familiale (parcours spatio-temporelle des enfants)

!! only for children not living in the house anymore, the children living in the house are already investigated in section 1!!

Quest-ce-que le nom, age, education et profession de votre enfants ? Ou est-ce qu'ils habitant maintenant? Pourquoi sont-ils se déplacés? (Suggestions: proximité de leur lieu de travail, à cause de raisons familiales, à cause de la disponibilité des transports en commun, ...)

CHILDREN	Nom	Age	Education	Profession (depuis quand)
1				
2.				
3.				
4.				
5.				
6.				

CHILDREN	Place habite	Pourqoi déplacé?
1		

2.		
3.		
4.		
5.		
6.		

5. Impression de l'évolution de la population

Pensez-vous que la population est en baisse et à la hausse dans ce village? et pourquoi?

6. Niveau d'équipement du village

Y at-il une école maternelle / école primaire / école secondaire dans le village? Vos enfants vont (encore aller) à cette école ? Combien d'enfants à l'école? Si aucune école: où ne vont les enfants? Est-ce qu'il y a des écoles a qui sont fermés au cours des dernières décennies? Ou y at-il de nouvelles écoles ouvertes dans les dernières décennies?

N°of école maternelle		Vos enfants à école ? Ou ?	
N°of école primaire		Travel time&distance vos enfants école?	
N°of école secondaire		Écoles fermé / ouverté last decennie?	

Combien d'épiceries, supérettes, supermarchés, boucheries et boulangeries sont dans le village? Ou faites vous vos courses? Quel est l'endroit où vous allez pour les éléments de base, comme la nourriture ...? Alles vous dans d'autres magasins? Quel types? Ou sont-il situés? Y at-il une évolution particulière du nombre de supérettes et supermarchés dans ce village les dernières décennies? Y at-il minimarkets ou des supermarchés fermés les dernières décennies, ou ouvert?

Combien de pharmacies(/ sont là dans le village + évolution? Combien de médecins sont là dans le village + évolution? Combien de boulangeries sont là dans le village? + evolution Combien de boucheries sont là dans le village? + Evolution? Comment est l'évolution du nombre de cafés/restaurants de la dernière décennie? Est-ce que le nombre de cafés en hausse ou en baisse? Savez-vous pourquoi? (Combien…….)

Y at-il une maison de retraite? Sont-ils de plus en plus / en déclin, combien de personnes vivent en elle? Quelle est la Distance à la plus proche maison de retraite?

N°of épiceries		Ou magasinent-vous?	
N°of supérettes		Travel time & distance supermarchés	
N°of supermarchés		Evolution magasin (all kind: boucherie boulangerie, supermarché …)	
N° of boulangeries		N° of restaurants / cafés + evolution	
N° of boucheries		N° of pharmacies + evolution	
N° of autre magasin		N° of médicins + evolution	
N° of églises		N° of suares, carrés	
N° of Paroisse Halls		N° of club de jeunes	
N° of club sportifs		Maison de retraite ? Travel Time & distance to closest MDR	

Y at-il un changement dans la disponibilité des transports en commun au cours des dernières décennies? Y at-il plus ou moins les bus maintenant dans le présent rapport au passé?

7 facteurs de migration contrôle: qu'est-ce que vous trouvez importantes dans ce village? Voici les facteurs importants pour vous? Ce qui rend le village un endroit agréable à vivre? Qu'est-ce qui pourrait être amélioré?

Est-il-important de vous d'avoir (des upermarchés et magasins près de chez vous, des transports public à proximité de votre maison, une grrande ville pas trop loin de l'endroit ou vous vivez, des cafés et restuarants, club de sports, ….)

	Pas important (0)	Un peu important (1)	Important (2)	Très important (3)
Nature (environment vert)				
Près de lieu de travail				
Près de large ville				
Cafés, restaurants dans le village				
Club sportifs dans le village				
Médicin dans le village Pharmacie dans le village				
Écolle maternelle& primaire dans le village				
Distance supermarché				
Distance école sécondaire				
Transport public dans l'environ				
Église dans village				
Prix de la construction / prix de louer maison				

Selon vous, quelle sont les évolutions les plus importantes au cours des dernières décennies dans ce village?

Appendix B: Variables used in correlation analysis

Variable	Unit	Data source	Comments
HP	Price of buying a house in euros/m²	Meilleurs Agents immobilier Data used from October 2016	Tricky, no history of HP available; HP of October 2016 on website of MeilleursAgents are used.
EP (normalized)	Available employment per inhabitant in and around a village	Insee.fr + own processing	Calculated with inverse distance weighting , normalized per inhabitant
DEP	Evolution of available employment	Insee.fr + own processing	
SL (normalized)	Between 0 and 1	Field Work data	Weighted SL: if people indicated some services were more important than others, they received a higher weight in the calculation of the (normalized) SL.
DSL	From -1 (-100%) to +1 (+100%)	Field work data	During field work people were asked about the evolution of the number of services in their village the last 40 years and the opening of new services and closure of services in the past.
DTH	In km	Field work data	Asked during field work
Distance to regional city	In km	Field work data Google Maps	Asked during field work for the 24 visited villages. For the other villages of the Marne department, Google Maps was used to obtain data for more villages. The DTC defined as the distance from the center of the village to the center of the nearest city.
DTS	In km	Field Work Data Google Maps	Asked during field work for the 24 visited villages. For the other villages of the Marne department, Google Maps was used to obtain data for more villages. The DTS defined as the distance from the center of the village to the nearest supermarket.
Population density	In inhabitants/km²	Insee.fr Arcmap analysis	
%employment agricultural sector	0-100%	Insee.fr	
%employment industrial sector	0-100%	Insee.fr	
%employment service sector	0-100%	Insee.fr	
Forest fraction	0-100%	Analysis of Corine Land Use maps	
Elevation of village (center of village)	Absolute elevation in meter above sea level (m)	Google Earth	
Soil fertility	Categorical analysis	Soil Maps library + Field Work data	

Appendix C: Additional information regarding statistical terms

Some important concepts to understand the figures are the confidence interval and the prediction interval. The confidence interval delimits the range of the true best-fit line in such a way that there is a 95% probability the true best-fit line lies within the confidence interval (Greenland et al., 2016). This implies that if a study is repeated 100 times, the chance that true best fitted line is within the confidence interval is 95% (Greenland et al., 2016). The prediction interval delimits the range of the individual data points. An individual data point has an 95% probability of falling within the domain of the prediction interval (Faraway, 2002).

A correlation coefficient (R^2) and the significance of the plotted regression line are indicated on the figures. The correlation coefficient indicates how much of the variance is explained by the regression line (Faraway, 2002). The p-value indicates the discrepancy from the null hypothesis. The null hypothesis assumes there is no relationship between the population evolution and the tested variable. A p-value of 0.05 implies there is a 5% chance we obtain the result by chance when there is no real link between the population evolution and the tested variable (Ferreira & Patino, 2015). The determination of the exact threshold is arbitrary, but often a p-value of 0.05 is chosen to decide whether to keep or reject the null hypothesis (Dahiru, 2008; Ferreira & Patino, 2015; Greenland et al., 2016).