Qualifier Report

Exploring the Spatial Diffusion and Knowledge Production Dynamics of the Clean Energy Sector and its Resilience to External Economic Shocks

Vidit Kundu

Faculty ITC, Department of Urban and Regional Planning
and Geo-Information Management

Examination Committee:
Prof. Dr. Karin Pfeffer, Dr. Debraj Roy, Prof. Dr. Michel L. Ehrenhard, Prof. Dr. Tatiana Filatova, Dr. Amineh Ghorbani

March 10, 2021

UNIVERSITY OF TWENTE.

Abstract

In order to achieve the established climate goals, clean energy industries are of high importance for policymakers with an additional promise of creating numerous new jobs. In the context of grand societal challenges, such as climate change and the associated transition toward renewable energies, the understanding of regional diversification (related or unrelated), and knowledge and its characteristics becomes particularly important for strategies aiming to develop new green paths or strengthening established green regimes. While the recent policy-related conceptual advancements ranging from related variety, local path dependencies, creation of regional advantages, and smart specialization are centred around knowledge exchange and learning processes, the operationalization of knowledge-related conceptual elements still remains a demanding task. Moreover, studies in economic geography usually focus on the role of the (local) supply side in the emergence and evolution of industries, emphasizing agglomeration externalities, windows of local opportunity, path dependence, and related variety, and have been less focused on developing hypotheses around multi-scalar drivers, vested interests, competing technologies, unrelated variety and the simultaneous analysis of demand and supply. This research aims to investigate the key factors responsible for the emergence and evolution of the various clean-energy sectors, specifically the wind, solar energy and electric vehicle sector, as well as the factors driving innovation within them. It aims to use network measures to propose a novel operationalization of knowledge at the aggregate level and operationalizes non-local knowledge sourcing through an antecedent-descendent dataset, thereby investigating the relevance of multi-scalar (local and non-local) knowledge dynamics in both the evolution of the various clean-energy sector as well as in their innovation dynamics. An improved understanding of knowledge and its characteristics – including its diffusion, creation, the qualitative definition of aggregate knowledge and evolution of knowledge networks within the various clean energy sectors is expected to positively influence innovation policies by helping policymakers shift their focus from correcting for market failures to more systematic problems. The research also intends to study the co-evolutionary dynamics between the local knowledge production and demand creation, as well as the competing technologies contesting the process of diversification. Finally, the research aims to investigate the resilience of regional knowledge production, and niche and regime formation dynamics, when exposed to an external (economic) shock and tries to simulate possible regional resilience scenarios. The study intends to make use of statistical analysis methods (including, OLS regression, Bayesian survival analysis, etc.), network analysis methods (SNA, SAOM, etc.) as well as an agent-based simulation model for the above-intended purposes.

Contents

1. Introduction

In the context of grand societal challenges, such as climate change and the associated transition toward renewable energies, the understanding of regional diversification (related or unrelated) and knowledge and its characteristics becomes particularly important for strategies aiming to develop new green paths or strengthening established green regimes. In order to achieve the established climate goals, clean energy industries are of high importance for policymakers with an additional promise of creating numerous new jobs (Burton et al., 2001). Even though knowledge exchange and learning processes form the heart of the recent policy-related conceptual advancements ranging from related variety (Frenken et al., 2007), local path dependencies (Martin & Sunley, 2006), creation of regional advantages (Asheim et al., 2011), and the idea of smart specialization (Morgan, 2015), the operationalization of knowledge-related conceptual elements still remains a demanding task. All these theoretical contributions are based on the recognition that innovation is a systemic phenomenon. An improved understanding of knowledge and its characteristics – including knowledge diffusion, creation of new knowledge, the qualitative definition of aggregate knowledge and evolution of knowledge networks within the clean energy sector is expected to positively influence innovation policies by helping policymakers shift their focus from correcting for market failures to more systematic problems. Moreover, the existing empirical literature has been less focused on developing hypotheses around multi-scalar drivers, vested interests and competing technologies contesting the diversification process, and the simultaneous analysis of demand and supply factors still remains rare. This study contributes to and connects the literature on evolutionary economic geography, technology transition, spatial diffusion, innovation systems, agglomeration theory and industry life-cycles by disentangling the relevance of multi-scalar knowledge dynamics, demand and supply, and the various constraining factors in the emergence of geographical clean-energy sector niches and their subsequent development into regimes.

Therefore, this research primarily tries to analyse the role of multi-scalar (local and non-local) knowledge dynamics in the emergence and the evolution of the various clean-energy sector, specifically the wind, solar and electric vehicles industry. First, it aims to use network measures to define knowledge qualitatively and quantitatively at both the aggregate (local) and the individual level. It then tries to operationalize non-local knowledge sourcing through an antecedent-descendent dataset and investigates the relevance of both. It does this not just for the focal sector, but also for the related sectors in the region. Second, besides analysing the role of knowledge in the emergence and evolution of the sector, the study also tries to investigate its role in demand creation, thereby trying to disentangle the co-evolutionary dynamics between demand and supply (here, multi-scalar knowledge production). Third, the study also investigates the role of knowledge (both aggregate and individual) in the innovation dynamics within a region, as established through patent filings and new combinations of knowledge elements. Fourth, it aims to study the driving factors behind knowledge network formation and evolution in order to be able to simulate aggregate regional knowledge production and innovation. Finally, the study aims to investigate the resilience of this regional knowledge production as well as the niche and regime formation dynamics in the case of an external (economic) shock. Below I unpack the various gaps that this research intends to address. First, I discuss the lack of understanding in network evolution dynamics. Then, I focus on the lack of qualitative and quantitative operationalization at the individual and the aggregate level, while also discussing the need for using agent-based models. Thereafter, I highlight the lack of attention that the demand-side, multi-scalar drivers, vested interests and competing technologies have received in the economic geography literature. Finally, I discuss how addressing the above could also help in predicting regional resilience when faced with external (economic) shocks.

Presently there is a considerably strong focus on inter-organisational cooperation which can be explained by the fact that "invention activity is far from being the outcome of isolated agents' efforts but that of interactive and collective processes" (Carayol & Roux, 2009, p. 2). In most industries, pronounced intensity of R&D cooperation and the emergence of innovation networks with dynamically changing compositions over time is observed (Kudic, 2014). At the same time, it is known that the economic actors' innovativeness is greatly affected by their strategic positioning in the network. Networks provide a natural infrastructure for knowledge exchange and markets by enabling information and knowledge flow (Podolny, 2001). While research on knowledge networks has increased over the years providing some understanding about how they emerge and evolve over time, there is still limited understanding of how space (Balland et al., 2013), other knowledge related factors such as the modularity of the knowledge-base and technological complexity, and network measures such as closeness and betweenness centrality mediate this process (Buchmann, 2015). Moreover, while a few studies have investigated the driving forces behind network formation in other sectors (Balland, 2012; Balland et al., 2013; Broekel & Boschma, 2012; Cassi & Plunket, 2010), empirical research in the renewable sectors is still sparse. Additionally, the current literature focuses on identifying mechanisms behind knowledge network evolution, taking into consideration different kinds of network properties (embeddedness, proximity, and status) (Balland et al., 2019; Broekel & Boschma, 2017; Cassi & Plunket, 2015). Little consideration has been given to both the knowledge production *across* technological fields and local and non-local knowledge flows (Esposito, 2020; Tsouri et al., 2020). Finally, very few studies have focused on the interaction effects between the different kinds of proximities (specifically, geographical, cognitive and social) at various spatial scales, and their role in knowledge network development and innovation performance (Broekel & Boschma, 2017; Cantner et al., 2015). This study aims to investigate the relevance of these factors, simultaneously with other factors, in the emergence and evolution of knowledge networks during the different life-cycle stages of an industry.

In order to further analyse innovation and new path development processes, it is thought to be necessary to move beyond relying on simple counts of knowledge inputs or outputs (e.g., patents, number of new products). Innovation policy-related conceptual advancements are based on the recognition that innovation is a systemic phenomenon. In line with Pyka et. al. (2019), this study argues that legitimations of policy interventions based on market failures need to be critically assessed and complemented by process-oriented reasoning that acknowledges the significance of individual and collective learning processes at higher aggregation levels. At the same time, analysing the consequences of interventions in complex adaptive innovation systems has proven to be challenging. Agent-based modeling "provides a powerful tool to account for the causes and consequences related to system dynamics, to analyse and compare system characteristics with a pre-specified benchmark, and to point to potential systemic misspecifications (Pyka & Fagiolo, 2007) and to bridge the gap between system failure categories and ex-ante policy evaluation" (Pyka et al., 2019, p. 1323). Recently, various agent-based models and simulations have emerged that try to capture the various dynamics of knowledge creation and diffusion in (innovation) networks. In previous models, knowledge has often been represented as a vector (or sometimes scalar) (eg., Cowan & Jonard, 2004; Mueller et al., 2017). However, the complex structure of knowledge generation and diffusion can be elusive if knowledge is treated as a number (or a vector of numbers) (Morone & Taylor, 2010). Arguably, the analysis of knowledge creation and diffusion processes may provide an incomplete account when representing knowledge as a simple quantifiable cumulative entity in terms of scalars or vectors (Schlaile et al., 2018). There are numerous examples where knowledge creation and diffusion involves more than 'stockpiling' additional pieces of knowledge, for example, by establishing complex relations. Therefore, in line with Morone and Taylor (2010) and Vermeulen and Pyka (2018, 2014b, 2014a), technological knowledge of an agent is perceived as 'units' that are ordered in directed graphs. While studies measure aggregate knowledge through the mean average knowledge stocks of all agents (Cassi & Zirulia, 2008; Morone et al., 2007; Morone & Taylor, 2004; Zhuang et al., 2011), the qualitative aspects of skills, the structure of the knowledge-base and the methods of analysis for studying such structures have largely remained unexplored (Yayavaram & Ahuja, 2008). Therefore, this study uses network analysis to investigate the structure of aggregate knowledge and network measures such as centrality, modularity and technological complexity to describe it.

Further, while a growing literature is focussing on the emergence and evolution of industries across space and time, the demand side has received relatively less attention. The economic geography literature usually focuses on the role of the (local) supply side in the emergence and evolution of industries, emphasizing agglomeration externalities, windows of local opportunity, path dependence, and related variety (Boschma & Wenting, 2007; Fornahl et al., 2012). Recently, the literature on technology transition has underlined demand processes as crucial for the emergence and expansion of industries and their products (Geels, 2004). In this literature, (market) niches are created with the help of local demand (in combination with local institutions) within which new industries can grow before facing the competitive forces of non-local markets (Schot & Geels, 2008). However, while there is a substantial empirical literature that assesses the role of regional supply-side factors for industries' emergence, empirical studies on the relevance of demand are less extensive (Bednarz & Broekel, 2020). Accordingly, it is still not well understood how local demand relates to the emergence and concentration of industries, and to what extent industries may themselves contribute to the activation and formation of local demand. At the same time, research in economic geography has recently been urged to adopt more multi-scalar perspectives on industrial path development (Hassink et al., 2019; Heiberg et al., 2020) as well as account for unrelated diversification (Boschma et al., 2017). Evolutionary economic geography (EEG) studies depict new regional industrial paths as outcomes of local structures and processes and underestimate non-local sources and influences (Hassink et al., 2019). Scholars in EEG have only recently started to systematically consider the importance of external (non-local) knowledge as a source of new path development (Boschma et al., 2017; Neffke et al., 2018; Trippl et al., 2018). However, further empirical studies are required to establish the relevance of external knowledge in the emergence and concentration of industries.

Another shortcoming of EEG is that it has thus far only delivered partial answers to the question of how industrial paths may be related (Hassink et al., 2019). EEG conceptualizations have primarily focused attention on the positive effects of relatedness between paths. There is very limited discussion of how several established paths located in a region may hinder each other or new paths in their development through competition over scarce assets and markets and other forms of negative path interdependencies (Hassink et al., 2019; Martin & Sunley, 2006; Steen & Hansen, 2018). Similarly, little attention is paid to the question of how new paths shape each other's evolution. More research is needed to reveal not only which constellation of factors (such as vested interests, or competition between established and emerging paths over scarce resources (e.g., Steen & Hansen, 2018) suppresses new path development, but also whether the relation between new industrial paths is competitive or complimentary. Therefore, besides trying to disentangle the role of multi-scalar knowledge in the emergence and concentration of industries, this study also tries to take into account other factors such as vested interests and competing technologies while studying the regional diversification process.

Finally, the study aims to investigate the resilience of the geographical configurations of niche formation and regimes when exposed to external (economic) shocks considering that regions, today, are more exposed to external shocks due to their increasing openness and interdependencies with the global economy (Toth et al., 2020). The ongoing coronavirus pandemic is the most recent evidence of this, which has left both the global and local value chains around the world severely shocked (Gereffi, 2020). The firm mortality resulting from the economic shock and firm entry through a greater impetus towards localizing global value chains is argued to have an impact on knowledge production at various scales, which might affect the geographical configurations of niche and regime formation and in turn the regional resilience. This study aims to investigate the above-outlined issues through the example of the clean energy sector, specifically wind, solar and electric vehicle industries.

2. Background

In the background section, I first review literature on the emergence and evolution of industries in general, and specifically as it pertains to the clean-energy sector in order to elucidate various factors associated with diversification and new path developments processes. Thereafter, considering that the study is centred around knowledge and its characteristics, I give an overview of the knowledge diffusion and innovation literature as well as provide a brief overview of the different models used to simulate knowledge diffusion and co-creation.

2.1 Emergence of Industries

The first strand of literature that focused on the spatial emergence of new industries was introduced by Scott, Storper and Walker (Scott & Storper, 1986; Walker & Storper, 1989) and came to be known as the **Window of Locational Opportunity (WLO**). According to the theory, new innovative agglomerations are able to form when impactful innovations erode the competitive advantages of incumbent regions (Boschma & Frenken, 2006; Esposito, 2020). Impactful ideas can travel to 'new' regions due to several idiosyncratic factors where they are used to make yet more ideas, thereby initiating local knowledge production. It was initially considered that because the resources and skills required for the new industry are generally diverse from the existing industries, their initial locations are distributed relatively arbitrarily (Boschma & Lambooy, 1999). However, lately it has been observed that the presence of related industries increases the likelihood of a new industry emerging in a region (Boschma et al., 2012, 2017). Rigby (2015) studied US metropolitan regions and found that technologies which were related to pre-existing technologies had a higher probability to emerge in that region (Boschma, 2017). Other studies have also replicated this finding (Boschma et al., 2013; Kogler, 2015). Evidence of this has also been found in the clean-energy sector (Bednarz & Broekel, 2020; Breul et al., 2015). This process of emergence and concentration of industries on the basis of regional related industries is known as **regional branching or related diversification** and manifests itself through diversification or spin-off mechanisms (Asheim et al., 2011; Boschma & Wenting, 2007). As Bednarz and Broekel (2020, p. 3) note, "both mechanisms, related diversification and spin-offs, have a strong regional dimension in that firms tend to establish new activities near to existing operations, and spin-offs tend to be located close to their parent company." Furthermore, spatially limited labour mobility and local knowledge spillovers within social networks, which also tend to be more intense locally, are other reasons that fuel regional diversification (Asheim et al., 2011). In the European Union policy circles, the process of regional diversification is increasingly regarded as crucial for smart specialization strategies aiming to develop new industrial paths in regions (Foray, 2014). **This research aims to investigate the key drivers responsible for the emergence of a new green industry in a region.**

As Boschma et al. (2017) note that while the process of related diversification has formed the focus of several recent studies in economic geography, the process of **unrelated diversification** still remains largely unexplored. A few studies have shown that unrelated activities are likely to develop in more knowledge-intensive regions in Europe as compared to knowledge-extensive regions (Xiao et al., 2018). Scholars have also indicated that non-proximate links facilitate unrelated diversification (Crespo et al., 2014), however, systematic evidence is still lacking. Therefore, in addition to investigating the role of local knowledge, **this research also aims to operationalise non-local knowledge sourcing and investigate its role in the emergence of new industries**, thereby also examining the process of unrelated diversification.

The other strand of literature that focusses on **technological transition** also provides inroads to address the above issues (Geels, 2002). It proposes a meso-level account of diversification, "as embodied in the creation of niches through systemic alignment that is attentive to both constraining and enabling factors internal and external to a region, and it addresses how niches for experimentation are mindfully created and scaled in the context of regimes through co-evolutionary processes of shielding, nurturing and empowerment…In this perspective, actors are essentially forced to adopt a bricolage mode of innovation while having to cope with vested interests and technological and cognitive lock-ins within established socio-technical regimes" (Boschma et al., 2017, p. 5). Therefore, the focus is more on how experimentation leads to change (Coenen & Truffer, 2012). A study done by Garud and Karnøe's (2003) on unrelated diversification, contrasted the emergence of the wind turbine industry in Denmark and the United States. It argued that while the US pathway represented a typical knowledge intensive strategy which tried to optimize for related variety, in Denmark, the wind industry developed through a bricolage process involving a broad set of local actors engaging in trial and error, and mobilization of financial, institutional and knowledge resources. This perspective highlights the importance of existing socio-technical regimes and various institutional and constraining factors in the processes of regional diversification. It also highlights the importance of establishing local technological niches, and consequently, the activation of local demand in the growth of an industry. Therefore, along the lines of Bednarz and Broekel (2020), **the research aims to explore the role of local supply-push (as understood through multi-scalar knowledge production) in local demand creation and the subsequent development of new industries, as well as the role of existing regimes and competing technologies in the process.**

2.2 Evolution of Industries

While new industries can emerge in some regions, not all of them benefit from the subsequent growth of industries as the concentration processes lead to agglomeration in only a few regions (Bednarz & Broekel, 2020). In evolutionary economic geography, this is explained through **spin-offs and agglomeration externalities mechanisms** (Klepper, 2006). Spinoffs processes refer to the process of founding new firms by former employees of incumbent firms in the same industry (Boschma & Wenting, 2007). The importance of spin-off in the growth and spatial concentration of industries has been demonstrated by a number of examples, such as the information and communication technology sector in Silicon Valley (Saxenian, 1994) and the United States automobile industry in Detroit (Klepper, 2002). As the likelihood of further spin-offs depends on the number of existing firms in a region, the mechanism can be considered to be a self-reinforcing process (Bednarz & Broekel, 2020). Besides spinoff dynamics, agglomeration externalities may affect the spatial evolution of an industry. These externalities arise from the co-localization of economic actors (Neffke et al., 2011) and are representative of the increasing returns in space (Krugman, 1991). Moreover, one can distinguish between **Jacobs and Marshall externalities**. The former refers to externalities arising from the co-localization of firms in different sectors and is otherwise known as urbanisation externalities. The latter relates to externalities arising from the agglomeration of economic actors in the same sector (Boschma & Wenting, 2007). As Bednarz and Broekel (2020) note, the discussion of these types of externalities goes beyond just considering the effect of regional conditions over industries, and also takes into account how industries themselves contribute to regional conditions. The concept of **related variety** stresses the role of related industries in concentration promoting externalities (Boschma et al., 2017; Boschma & Wenting, 2007).

When considering the industry's lifecycle, it is expected that the relevance of the different processes and factors like spinoffs, branching, agglomeration externalities, related variety etc. will be different in different stages of the lifecycle of an industry. Spinoff mechanisms are expected to be less dominant in the early stages of the lifecycle of an industry due to low numbers of the parent companies (Boschma & Wenting, 2007). The same is considered to be true for agglomeration externalities that are expected to be more relevant in later phases. New industries tend to benefit from new knowledge, which old industries are not suited to provide (Boschma & Lambooy, 1999). Several empirical studies confirm these processes (Balland et al., 2013; Neffke et al., 2011). For instance, Neffke et al. (2011) show that where young industries benefit from Jacobian externalities in the initial phase, Marshallian externalities are more beneficial for mature industries. Studies find evidence of this in the wind and solar energy sector (Bednarz & Broekel, 2020; Breul et al., 2015).

Further, Binz and Truffer (2017) highlight the relevance of the typology of knowledge in the spatial diffusion of a sector. Considering that the the wind industry is built upon experience-based skills that are hard to replicate and unlikely to diffuse in space, it is characterized as a spatially sticky innovation system. The spatial diffusion of these industries, in the growth phase, is likely to remain the same as the initial phase. On the other hand, other industries, such as the photovoltaic industry that majorly employ scientific-technical skills can be considered relatively more *footloose* as the regions diversifying into these industries are more prone to losing the agglomeration when faced with competition from other regions. Moreover, as the role of scientific-technical knowledge has increased in the wind (as well as the auto) industry over the years, it is also losing its spatially sticky nature. However, empirical evidence of this is still limited. Therefore, **this research aims to not only investigate the key drivers responsible for the evolution of a clean energy sector in a region, but also tries to analyse the role of increasing complexity of knowledge in the changing spatial diffusion patterns of the various clean-energy sectors observed over the years.**

2.3 Knowledge and Innovation

2.3.1 Overview

Given that the acquisition and creation of (new) knowledge gives considerable competitive advantages, there is a keen interest of firms and policy makers on how to foster the creation and diffusion of (new) knowledge. A firm can develop its knowledge base either internally through R&D activities or externally by mobilizing knowledge resources through interaction with other actors (Malerba & Torrisi, 1992). However, due to an ever increasing complexity of technology, the knowledge base of a single firm is rarely sufficient for innovation. Innovative firms need access to a greater quantum of knowledge, which they frequently access through collaborating with other firms. A pronounced intensity of R&D cooperation is observed in most industries along with the emergence of innovation networks with dynamically changing compositions over time (Kudic, 2014). At the same time, it is known that the economic actors' innovativeness is greatly affected by their strategic positioning in the network. Networks provide an infrastructure for knowledge exchange and markets by enabling information and knowledge flow (Podolny, 2001).

Information and knowledge can be considered distinct concepts, while the former is of factual nature, the latter involves cognition and learning (David, 2002). The process of learning could be understood as a process under which information is acquired and assimilated with previously acquired information and subsequently generalisations and correlations made (Howells, 2002; Saviotti, 1999). Consequently, a necessary precondition for learning is the presence of related acquired knowledge. Similar to drawing a distinction between information and knowledge it is also important to distinguish between tacit and codified knowledge. Tacit knowledge is also referred to as the knowledge that is 'person-embodied, context-dependent, spatially sticky and socially accessible only through direct physical interactions' (Morgan, 2004). This definition does not considerably differ from the original idea of tacit knowledge as introduced by Polanyi (1967), who describes it as the component of knowledge which is distinct from, but complementary to conscious cognitive processes. Codification on the other hand involves standardization. According to Saviotti (1998), codification can be understood as the gradual convergence of users around common concepts, definitions, contents and theories. Both tacit and codified knowledge are contextual, as at different periods of time a different proportion of knowledge will be either tacit or codified. As new knowledge is understood better over time, it gets more codified (Cowan & Foray, 1997).

The tacit component of knowledge is the main reason behind the geographical concentration of innovative activities, primarily due to its requirement of interpersonal interactions in the learning process (Meder, 2008). Certain sectors, such as the early wind industry, biogas, or luxury watchmaking are majorly built upon experience-based skills that are hard to replicate (Binz & Truffer, 2017). This makes them spatially sticky and restricts their diffusion in space throughout their lifecycle. Thomas Allen (1977) proposed the 'Allen Curve' which revealed a strong negative correlation between the physical distance amongst the employees and their frequency of interactions and collaborations. This reveals the importance of co-location in knowledge-exchange processes.

Furthermore, innovation is increasingly dependent on interactions and knowledge flows between firms, universities, research institutes and public agencies, and much of these interactions occur locally (Gertler, 2003). This explains the increasing geographic concentration of innovation activities in a world of global markets. The effect is strikingly pronounced for knowledge intensive activities (Balland et al., 2020).

Diffusion literature has recently focussed considerable attention on innovation networks and how knowledge diffusion within (and outside of) these networks (Morone & Taylor, 2010; Schlaile et al., 2018). As Schlaile et al. (2018) describe, the literature has tried to answer four key questions in this regard. These are – a) What diffuses? b) How does it diffuse? c) Where does it diffuse? d) What are the effects (or performance) of the diffusion process and how are they measured?

Regarding question, a), most studies focus on analysing knowledge in simplified vector and scalar quantities. Arguably, the analysis of knowledge creation and diffusion processes may provide an incomplete account when representing knowledge as a simple quantifiable cumulative entity in terms of scalars or vectors (Schlaile et al., 2018). Only a few studies represent knowledge elements as network structures (Cassi & Zirulia, 2008; Cowan & Jonard, 2004; Morone & Taylor, 2004; Schlaile et al., 2018). Concerning question b), how it diffuses, Morone and Taylor (2010) suggest decomposing knowledge diffusion into three sub categories that include knowledge spillover, knowledge transfer and knowledge integration. The first two denote similar processes, the only difference being that while spillovers could be considered unintended, knowledge transfer is an intended process. Another sub category allows for integration of knowledge without acquiring it. This is termed as knowledge integration and is a faster way of collaboration that is well suited to the rapid innovation landscape. Regarding c), where it diffuses, studies mostly focus on archetypical network structures, such as small-world structures, random structures or regular structures. While static networks have been studied more (Cassi & Zirulia, 2008; Cowan & Jonard, 2004), a few studies also look at network structures evolving over time (Morone et al., 2007; Zhuang et al., 2011). Studies mostly measure d) through the mean average knowledge stocks of all agents (Cassi & Zirulia, 2008; Morone et al., 2007; Morone & Taylor, 2004; Zhuang et al., 2011) Finally, knowledge dynamics could be thought of as outcomes based on the aggregation of flows between individual units (Morone & Taylor, 2010). Specific flows could be used for explaining innovation levels for the larger regional economy.

The study intends to employ a network-of-networks approach, that means each agent (e.g., firm or individual) in an innovation network itself contains a network of knowledge units or skills (Schlaile et al., 2018). Knowledge is defined at two levels. While one is at the level of an individual agent (here, an organization), the other is at an aggregate regional level. The next section looks at the literature on knowledge network formation in more detail.

2.3.2 Knowledge Networks and Proximity

The literature on knowledge networks has majorly focussed on the driving factors behind their evolution considering properties at the nodal, relational and structural level (Balland et al., 2013, 2019; Cassi & Plunket, 2015). Dyadic proximity of actors, their embeddedness in the network, and actor level characteristics are the main mechanisms identified behind knowledge network evolution (Balland et al., 2016; Giuliani, 2013).

Network theory suggests that in addition to the attributes of the individual actors involved in knowledge exchange, their relational characteristics are equally important for network evolution. Correspondingly, the mutual similarity between two actors would influence the probability of link formation or dissolution. In economic geography, this is conveyed through the concept of proximity. In brief, the concept states that proximities in different dimensions could help in fostering innovative activities and interactive learning (Roesler & Broekel, 2017). Inspired by the French School of Proximity Dynamics, Boschma (2005) developed a concept of five different proximities that was seen useful in understanding interactive learning and innovation. These include geographical, cognitive, organizational, social and institutional proximity. '**Cognitive proximity**' or '**technological proximity**' refers to the degree of similarity between the knowledge bases of two actors. An effective knowledge transfer between actors requires a complementary absorptive capacity to identify, interpret and exploit the knowledge that is being transferred (Cohen & Levinthal, 1990). Hence, the similarity in knowledge bases facilitates knowledge transfer and also serves as an incentive for actors to collaborate (Meder, 2008). However, it has been observed that certain cognitive distance is requisite to avoid technological lock-in that might arise due to the homogeneity of the knowledge involved (Broekel & Boschma, 2012). Therefore, much like other proximities, cognitive proximity not only facilitates knowledge exchange but also specifies the learning potential for the creation of novel ideas (Broekel & Boschma, 2017). '**Geographical proximity**' could be understood as the spatial or physical distance between actors. It has been seen to enhance both the collaboration probability (Feldman & Florida, 1994; Meder, 2008) and the outcome of the collaboration in terms of learning and innovation (Audretsch & Feldman, 1996; Cantner et al., 2015; Jaffe et al., 1993). '**Social proximity**' in terms of collaboration could be interpreted as learning effects originating due to the embeddedness of actors within a social network. Social proximity could lead to repetitive cooperation between the same set of actors (Boschma, 2005). '**Institutional proximity**' is associated with the institutional framework at the macro scale (Boschma, 2005). Organisations belonging to the same institutional setup share similar routines, practices rules or laws (Edquist & Johnson, 1997). These routines could include similar research procedures, standards or administrative structures and allow for easier knowledge transfer. Finally, '**organisational proximity**' could be understood as the degree to which organisations share the same institutions (Boschma, 2005). Therefore, understood over a continuum, low organisational proximity would mean independent actors with no ties, whereas high organisational proximity would mean actors organised hierarchically within the same institution with strong ties with one another.

As Broekel and Boschma (2017) point out, the benefits that an actor can gain from certain links is related to optimum levels of proximity in various dimensions. Certain studies (Broekel & Boschma, 2012; Nooteboom et al., 2007) provide evidence that in terms of innovation output, an optimum level of cognitive proximity might exist. That means both greater technological similarity or lower technological similarity between actors could reduce innovation performance. Boschma & Frenken (2010) make a distinction between cognitive proximity on the one hand and the rest of the proximities on the other hand. For all proximities, except cognitive proximity, the 'optimum level' refers to a mix of both low and high proximity links. For instance, in the case of geographical proximity that would mean a balance of local and non-local links and not an optimal geographical distance between two agents. This applies to other proximities as well. However, in the case of cognitive proximity, an optimal level of proximity needs to exist for each link. Empirically, this means to refrain from calculating an *average* technological similarity (cognitive distance) of a focal firm to its knowledge exchange partners, which many studies do (Broekel & Boschma, 2012; Nooteboom et al., 2007), but instead adopt a clustering approach (Broekel & Boschma, 2017).

Boschma (2005) also points out that proximity types could relate to each other as either substitutes or complements in establishing links and ensuring their success. In the case of substitutes, being proximate in just one dimension is sufficient to form a link, without the need to be proximate in other dimensions (Boschma & Frenken, 2010). Some studies provide evidence for this, including studies showing geographical proximity fostering link formation among cognitively distant researchers (Singh, 2005). Other studies show that geographical proximity helps overcome institutional distance, in that university-firm relationships are more likely when the university and firm are geographically proximate, while university-university relationships could still occur at a greater geographical distance (Ponds et al., 2007; Roesler & Broekel, 2017). However, recent studies point to there being a more complex relationship between geographical and cognitive proximity, where actor attributes influence whether the relationship would be substitutive or complementary (Broekel & Boschma, 2017). **This study further tries to investigate the relationship between geographical proximity and cognitive proximity within the clean energy sector.**

While research on knowledge networks has increased over the years providing some understanding about how they emerge and evolve over time, there is still limited understanding of how space (Balland et al., 2013) and qualitative aspects of knowledge elements mediate this process. Moreover, while a few studies have investigated the driving forces behind network formation in other sectors (Balland, 2012; Balland et al., 2013; Broekel & Boschma, 2012; Cassi & Plunket, 2010), applied research in the renewable sectors is still sparse. Moreover, while the current literature focuses on identifying mechanisms behind knowledge network evolution, taking into consideration different kinds of network properties (embeddedness, proximity, and status) (Balland et al., 2019; Broekel & Boschma, 2017; Cassi & Plunket, 2015), little consideration has been given to both the knowledge production *across* technological fields and the changing typology of knowledge. Additionally, very few studies have focused on the interaction effects between the different kinds of proximities (specifically, geographical, cognitive and social) at various spatial scales, and their role in knowledge network development and innovation performance (Broekel & Boschma, 2017; Cantner et al., 2015). **This study aims to investigate the relevance of these factors, simultaneously with other factors, in the emergence and evolution of knowledge networks during the different life-cycle stages of an industry.** Finally, it is understood that the importance of knowledge production in the development of an industry varies across sectors and timescales. Thereby, one is more likely to see knowledge networks spanning across multiple sectors specifically within knowledge intensive sectors, such as solar energy sector, due to a greater relevance of related knowledge in innovation. In this study, it means that there is a greater likelihood for related knowledge to be more crucial in the growth phase of the solar energy sector as compared to the growth phase of the early wind energy sector.

2.3.3 Modeling Knowledge and its Diffusion

Most initial models focussed at knowledge diffusion *of* innovation and not the knowledge sharing *to* innovate. **Neoclassical economic models** worked under the assumption of perfect information. However, empirical evidence shows that knowledge diffusion is a geographically localised process and only with time does the information about a new technological development spreads over a geographically wider region (Jaffe et al., 1993). Moreover, the models not only conflated diffused information with diffused knowledge, they also assumed that information spread was analogous to adoption of a new technology (Slicher van Bath, 1963). However, studies found that there is a lag of several years between leaning about a new technology and adopting it. Another class of diffusion models used were known as the **'epidemic models'**. These could broadly be subdivided into broadcasting and word-of-mouth models. The former referred to the models within which the knowledge about a new technology reached everyone through the same medium. On the other hand, the second subclass referred to the models in which diffusion happened through personal interactions. While the first subclass traced an exponential curve (with a negative exponent), the second was relatively more realistic which charts an S-shaped logistic curve (Geroski, 2000). However, infection probability in both cases were exogenously defined.

Subsequently, **game theoretical models** have been developed to study innovation diffusion, taking into account a greater level of complexity. Authors have introduced heterogeneous agent populations and technologies along with neighbourhood learning rules to investigate the aggregate diffusion (Ellison & Fudenberg, 1993, 1995) and study the relationship between the neighbourhood structure and the adoption of new technologies (Bala & Goyal, 1998). Other authors have used evolutionary game theory to examine local interaction games within which pay off depends on the neighbour's action (Morris, 2000). Although these models overcome certain limitations of epidemic models (by including heterogeneous agents, social network and simultaneous innovations), yet all of these models still reduce learning to a binary mechanism (of either knowing or not knowing about a new technology) and conflate knowledge with information.

Agent based simulation models represent the new class of models that have been used to capture the complexity involved in knowledge diffusion processes. They move beyond the hyper-rationality, representative agents and equilibrium of the previous models. In contrast to previous models, the search for an innovation is open ended taking into account that what is being searched for is unknown and only discovered with time (Morone & Taylor, 2010). They are also able to incorporate social networks over which interactions take place (Fagiolo et al., 2006). As Axelrod (1997) lays out, agent-based models can not only be useful to understand individual behaviour, but also aggregate behaviour and how large-scale outcome is a product of many individual-level interactions.

One of the initial influential model was proposed by Cowan and Jonard (2004). In this model, agents were arranged in one dimensional space and knowledge is operationalized as a vector quantity. The main goal of the exercise was to understand how aggregate knowledge levels increase as agents interact over time and how this growth is affected by the underlying structure of the network. The small world network turned out to be the network with the highest aggregate knowledge. However, much like other initial models, the network so considered was static in nature and exogenously defined. Along with this, agents were assumed to operate with a complete information of their neighbours and did not engage in open ended learning. Later models tried to correct for the shortcomings of this model. Cassi and Zirulia (2008) introduced cost-benefit assessments, while Morone and Taylor (2010) worked with a more complex network definition of knowledge. However, empirical calibrations have lagged behind. Table 3.1 tries to list a few agent-based knowledge diffusion models with varying operationalization and diffusion of knowledge, as well as varying underlying network structures over which diffusion takes place.

Table 3.1: Characteristics of different agent-based knowledge diffusion models

Models	Knowledge Definition	Knowledge Flow	Agents	Underlying Graph // Spatial Unit	Simulation/ Validation
Cowan and Jonard (2004)	Scalar	Bilateral Barter Exchange	Heterogeneous in innovation, & absorption	Random, Small World, Connected // Undefined	Simulation
Cassi and Zirulia (2008)	Scalar	Bilateral Barter Exchange	Heterogeneous in innovation, absorption & learning	Random, Small World, Connected // Undefined	Simulation
Morone and Taylor (2010)	Vector (Types of Skills), Branch Network	Integration/ Diffusion	Heterogeneous in the types & the number of skills	Undefined, Emergent // Regional	Simulation, Validation

Pyka et al. – SKIN model (Pyka et al., 2009)	Vector (Type of Knowledge – Scientific/ Technical)	Diffusion	Heterogeneous in the types & the number of skills, & capital endowment	Undefined, Emergent // Regional	Simulation, Validation
Vermeulen and Pyka (2018)	Vector (Types of Skills), Branch Network	Diffusion	Heterogeneous in the types & the number of skills, & capital endowment	Six different configurations defined // Regional, Inter-Regional	Simulation, Validation
Frenken (2006)	Vector	Diffusion	Heterogeneous in the types & the number of skills, and interactions within skills	Small World // Regional	Simulation
Schlaile et al. (2018)	Vector (Types of Skills), Network	Diffusion	Heterogeneous in the types & the number of skills	Undefined, Emergent // Regional	Simulation

2.4 Resilience

Finally, the study aims to investigate the resilience of the regional innovation system as well as the path development dynamics under the exposure of an external (economic) shock. Resilience can be understood as the ability of a system to cope with shocks (Tsouri & Pegoretti, 2020). Specifically, as Smith e. al. (2011, p. 88) put it, network resilience is "the ability of a network to defend against and maintain an acceptable level of service in the presence of challenges". The literature on the geography of innovation and knowledge networks has not extensively focussed on the resilience of networks (Boschma, 2015) and knowledge production. Similarly, the role of knowledge networks has attracted little attention so far in the literature on regional resilience (Tsouri & Pegoretti, 2020). However, both the structure of regional knowledge networks as well as the regional knowledge production may not only be important because of their effect on the ability of regions to adapt to shocks, but also because they may affect the ability of the regions to create new path developments (Boschma, 2015). Moreover, as Tsouri and Pegoretti (2020) note, regional knowledge network resilience and regional resilience are not synonymous, even though one may affect the other. The previously elaborated analysis in this study that tries to disentangle the role of local and related knowledge in the development of an industry (supply-side) works towards establishing a link between the two. Regional economies that heavily depend on knowledge intensive industries may be more vulnerable to shocks in case their regional innovation system is not resilient enough. That can happen in the case of removal of specific nodes or linkages when exposed to external (economic) shocks which may cause instability or discontinuities to the regional knowledge network.

3. Research Design

The main research gaps identified in the literature are summarized below.

Table 4.1: Summary of the research gaps identified

Research Gaps
1. Limited understanding of the role of multi-scalar knowledge dynamics and local demand in the emergence and evolution of the various clean energy sectors
2. Inadequate comprehension of co-evolutionary dynamics between local demand and local knowledge production in the development of the various clean-energy sectors
3. Lack of understanding of the role of network measures such as modularity, centrality and technological complexity of knowledge elements in the evolution of the various clean energy sectors and regional innovation
4. Limited empirical research on knowledge network evolution within the various clean energy sectors, and the role of geography in it
5. Lack of operationalization of aggregate and individual knowledge production that takes into account both the qualitative and quantitative aspects of the skills involved, and consequently limited agent based simulation models for individual and aggregate regional knowledge production indicating the propensity for innovation of different regions
6. Limited understanding of the resilience of the knowledge production within the various clean-energy sectors when exposed to an (external) economic shock and subsequent impact on the niche and regime formation dynamics

These gaps are sought to be addressed through answering three major questions as described below. Question 1 and 3 are conceived to form the subject of a journal article each, whereas question 2 is anticipated to form a focus of two journal articles.

3.1 Research Objective and Questions

Main Objective: To explore the spatial diffusion and knowledge production dynamics of the clean energy sector and its resilience to external economic shocks.

To achieve this general objective, the following research questions have been formulated.

1. What are the key factors responsible for the spatial emergence & concentration of the clean energy sector?

2. What are the key driving factors behind the formation & evolution of knowledge networks within the clean energy sector?

3. How do external (economic) shocks affect the knowledge production and spatial diffusion dynamics of the clean energy sector and in turn the resilience of a region?

A further operationalization of the research questions, guiding sub-questions and the methods that would be used to answer them can be found in section 3.4.

3.2 Research Approach and Conceptual Framework

Along the lines of three main questions, the research is divided into three major phases. The first phase explores the key factors responsible for the emergence and concentration of the clean energy sector. The second phase looks at the factors responsible for the emergence and evolution of the knowledge networks and simulates aggregate knowledge production within the clean energy sector at various scales. Finally, the third phase analyses the resilience of knowledge production and geographical niche and regime formation dynamics of the clean energy sector when exposed to an external (economic) shock. Moreover, it looks at the role that the clean energy sector plays in regional resilience and the cascading effects of firm failure within the focal sector on related industries.

To empirically investigate the above, the example of both the Dutch as well as the European clean energy sector is considered. Considering that the knowledge production and evolution of industries is studied through datasets available at the European scale (PATSTAT and Bureau Van Dijk), it is possible to study network measures of knowledge elements, their interdependencies and their role in above dynamics at a larger scale to situate findings obtained at a national scale in a wider context. The research aims to study the dynamics within the wind, solar energy and electric vehicle sectors at the NUTS 3 spatial scale. This is motivated by a number of reasons. First, in order to achieve the established climate goals, clean energy industries are of high importance for policymakers with an additional promise of creating numerous new jobs (Burton et al., 2001). Second, the electric vehicle, wind, and solar sectors in their modern form are only a few decades old, which increases the availability of data for the initial stages of the industry. Moreover, one can source the geographic location of wind turbines, solar panels and EV chargers, thereby allowing for approximating the geographic (and temporal) distribution of demand. Third, the reason to choose three different sectors is to be able to compare the spatial diffusion and knowledge production dynamics amongst sectors with different knowledge typologies, life-cycles, geographic scale of supply and demand dynamics, and natural resources requirements. Finally, the choice of these three sectors is motivated by their positioning in three different quadrants within the 'innovation-valuation framework' (Binz & Truffer, 2017), as shown below, thereby covering the spectrum of clean-tech industries with varied innovation and valuation dynamics.

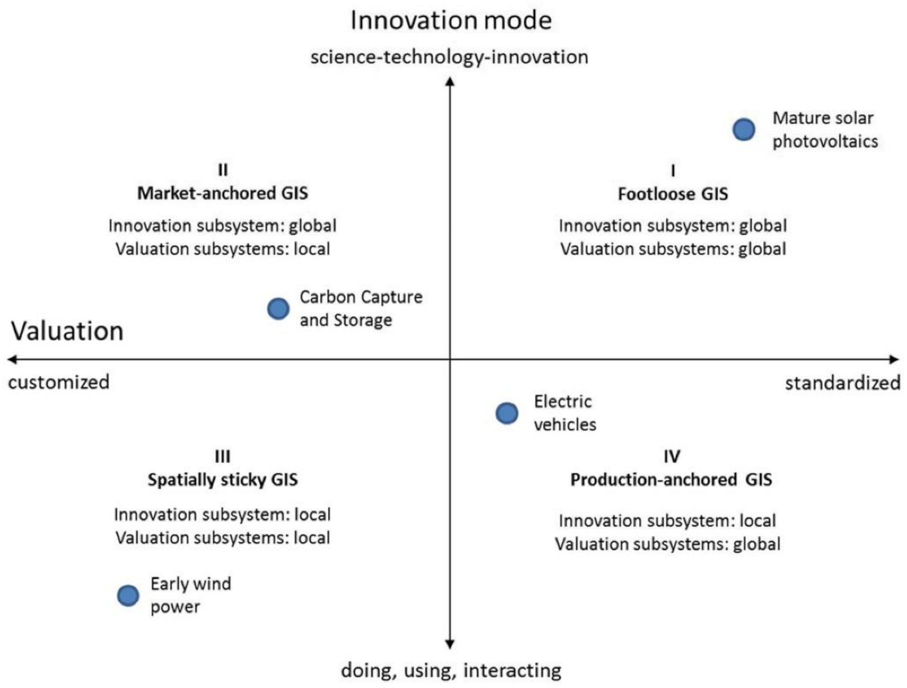
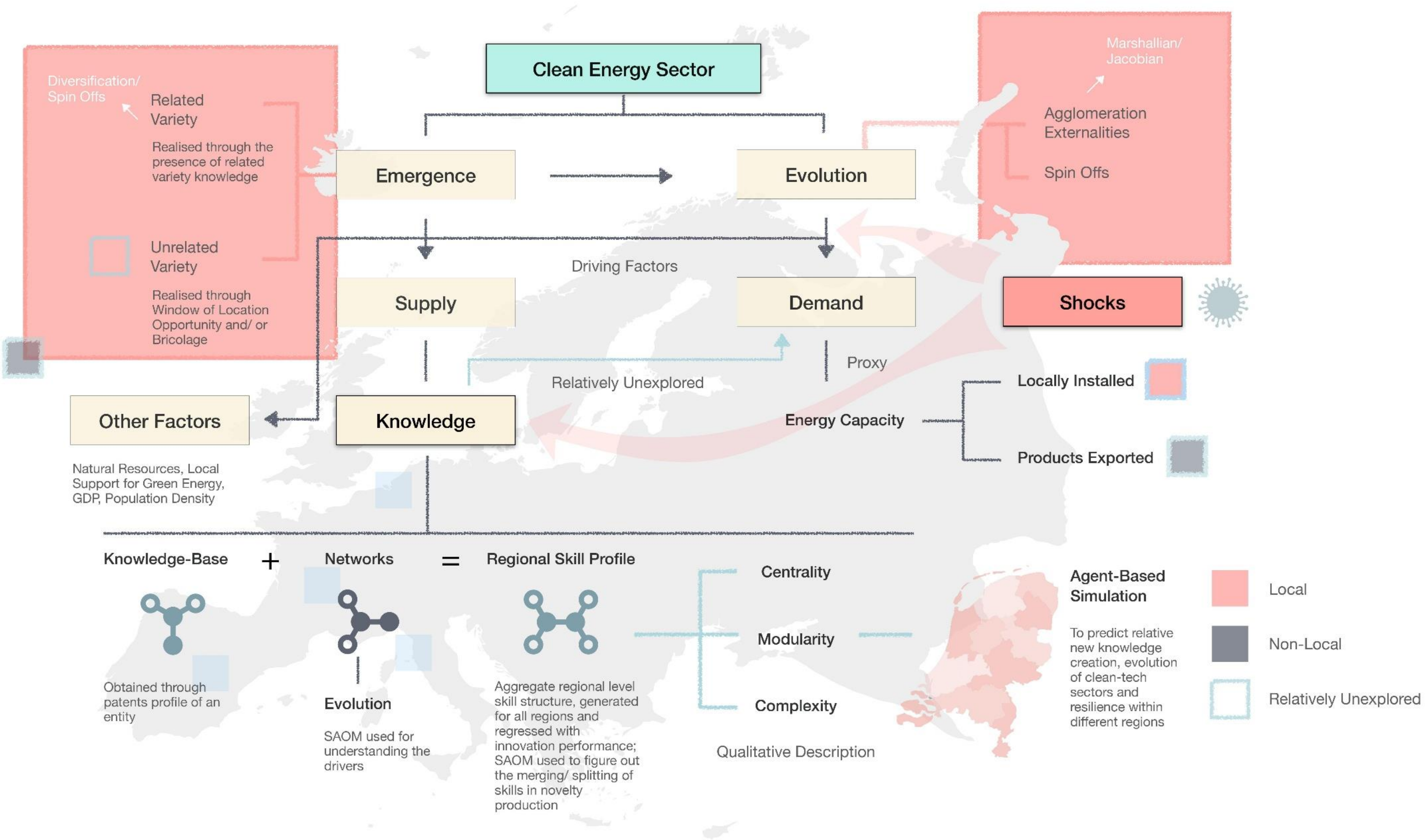


Figure: Positioning of different clean-tech industries in the innovation valuation framework. Source: (Binz & Truffer, 2017)

The conceptual framework is presented ahead. The framework illustrates how knowledge dynamics fall at the centre of regional diversification, agglomeration, innovation and resilience dynamics. The empirical variables and the methods are described further. This investigation, while centred on knowledge dynamics tries to answer questions regarding the success or failure of a region in terms of attracting renewable industries, advancing innovation, and withstanding shocks.



Clean Energy Sector
Diversification/ Spin Offs
Related Variety
Realised through the presence of related variety knowledge
Unrelated Variety
Realised through Window of Location Opportunity and/ or Bricolage
Emergence
Evolution
Marshallian/ Jacobian
Agglomeration Externalities
Spin Offs
Driving Factors
Supply
Demand
Shocks
Proxy
Relatively Unexplored
Locally Installed
Energy Capacity
Products Exported
Other Factors
Knowledge
Natural Resources, Local Support for Green Energy, GDP, Population Density
Knowledge-Base
+
Networks
=
Regional Skill Profile
Centrality
Modularity
Complexity
Obtained through patents profile of an entity
Evolution
SAOM used for understanding the drivers
Aggregate regional level skill structure, generated for all regions and regressed with innovation performance; SAOM used to figure out the merging/ splitting of skills in novelty production
Qualitative Description
Agent-Based Simulation
To predict relative new knowledge creation, evolution of clean-tech sectors and resilience within different regions
Local
Non-Local
Relatively Unexplored

3.3 Phases

3.3.1 Phase 1: Emergence and Concentration

The regional demand for wind, solar energy and EV industry is approximated through the time of activation, place of deployment and capacity of wind turbines, solar panels and EV charging stations. For this purpose, various databases are used. On the supply side, a detailed web research as well as including the AMADEUS database from Bureau Van Dijk is used to establish all the firms active in the different clean energy sectors in the Netherlands between 1970 and 2020 as well as their place and date of origination. Other datasets on renewable energy are collected from the different provinces of Netherlands.

In order to disentangle the co-evolutionary dynamics between the role of local demand in the emergence of an industry and the industry's role in developing local demand (supply-push), an empirical approach methodology similar to one used by Bednarz and Broekel (2020) in their study of the German wind industry, will be employed. **First Panel/ Turbine/ Charger** will be taken as a dependent variable and the time taken for the first solar panel, wind turbine, or an electric vehicle charger to be installed in region *i* will be modeled. The explanatory variables include **Manufacturers at t-5** that will take into account existing manufacturers in a radius of 25 kilometres with a time lag of 5 years given that a few years are required for the installation. The existing solar panels and wind turbines will also be taken into account with a time lag of 5 years (**Panels & Turbines at t-5**). Moreover, the effect of the parameters **Local Knowledge**, **Non-Local Knowledge** as wells as **Related Knowledge** will also be analysed. Unlike Bednarz and Broekel, this study will also take into account manufacturers and renewable energy capacity in related industries and investigates their role not only in the development of demand, but also in knowledge production dynamics.

On the other hand, the effect of local demand (demand-pull) on the probability of manufacturers emerging in a region *i* will also be analysed. **First Manufacturer** captures the founding of a new company in region *i*. The explanatory variable considered include **Panels & Turbines** as well as **Panels & Turbines at t+5**. The former is considered to capture the demand approximately 5 years ago, whereas the latter is assumed to capture the present demand due to prolonged installation periods. **Manufacturers** will be used to study the role of existing manufacturers in the emergence of a new manufacturer. This is to understand the importance of spatial clustering in the firm's emergence. As above, the role of **Local Knowledge Sourcing**, **Non-Local Knowledge Sourcing** as wells as **Related Knowledge** will also be analysed. Along the lines of Neffke et al. (2011) co-occurrence frequencies of 4-digit IPC classes on patents will be used to estimate cosine similarity between IPC classes which is used to obtain a measure of technological relatedness of each IPC class pair. Thereafter, Revealed Technological Advantage (RTA) for each IPC class and region will be calculated. Both matrices will then be multiplied with each other to obtain the aggregated relatedness coefficient for the focal classes in each region. Furthermore, other knowledge related explanatory variables will also be considered. These are described further in section 3.5.

Other factors might contribute to the diffusion of the clean-energy sector, which do not form a focus of this study. These will be assimilated through the use of control variables. **Solar/ Wind** will capture the average solar irradiance or wind velocity in region *i*, that might make a region ideal for locating solar panels or wind turbines. **Green Support** will capture the political inclination towards renewable energy. It is assumed that this variable would be less critical with the passage of time and with an increased consensus towards sustainable sources of energy. Finally, **Regional GDP** and **Population Density** of the region will also be controlled for.

Variable	Operationalization	Source
First Panel/ Turbine/ Charger	Time (t) of first installation of Solar panel/ wind turbine/ EV charger in the region i (NUTS 3)	Web Research/ Provincial Database
First Manufacturer	Time (t) of entering of the first Solar panel/ wind turbine/ EV charger manufacturer the market (NUTS 3)	Bureau Van Dijk/ Web Research/ Provincial Database
Panels & Turbines	Existing solar panels and wind turbines (including their capacity where possible) within 25km radius of the site of installation at time t	OpenStreetMap/ Web Research
Chargers	Existing EV chargers (including their capacity where possible) within 25km radius of the site of installation at time t	OpenStreetMap/ Web Research
Panels & Turbines at t-5	Existing solar panels and wind turbines within 25km, 50km radius at time t-5	OpenStreetMap/ Web Research
Chargers at t-5	Existing EV chargers within 25km, 50km radius at time t-5	OpenStreetMap/ Web Research
Panels & Turbines at t+5	Existing solar panels and wind turbines within 25km, 50km radius at time t+5	OpenStreetMap/ Web Research
Chargers at t+5	Existing EV chargers within 25km, 50km radius at time t+5	
Manufacturers	Existing manufacturers in NUTS 3 region i at time t	Bureau Van Dijk/ Web Research/ Provincial Database
Manufacturers at t-5	Existing manufacturers within 25km, 50km radius at time t-5	
Distance M-P/T/C t+5	Distance between new manufacturer and future panels/ turbines/ chargers	-
Distance P/T/C- M	Distance between new panel/ turbine/ charger and existing manufacturers	-
Related Knowledge	Co-occurrence of IPC classes in a patent and revealed comparative advantage for region i at time e	PATSTAT
Local Knowledge Sourcing	Number of patents whose antecedents lie within the focal NUTS 3 region at time t-3	

Non Local Knowledge Sourcing	Number of patents whose antecedents lie beyond the focal NUTS 3 region at time t-3	Dataset of descendent/ antecedent prepared from PATSAT
Solar/ Wind	Average solar irradiation/ wind speed in region i	Various Websites
Green Support	Percentage of green votes in the region i at time t	Manifesto Project Dataset
Regional GDP	GDP of region i at time t	Eurostat/ Provincial Database
Population Density	Population density of region i at time t	Eurostat/ Provincial Database

The founding of a firm in a location is modeled as an event in time, which is related to the prior emergence of other factors at another time and location. A Bayesian survival model is used to identify the determinants influencing the occurrence of this event. Survival models, unlike standard regression models, consider censuring (if the event occurs beforehand or after the observational period) which if not taken into account may skew results in a longitudinal data analysis.

In order to operationalize local and non-local knowledge sourcing, along the lines of Christopher R. Esposito (2020), this study makes use of patent classification codes to infer the flow of technological knowledge over space. The components based on the classification codes are used to define a knowledge set which is employed by a patent. These knowledge units are used to link patents to their parents (antecedents). The patent(s), amongst all the ones registered before the focal patent, with the maximum overlap of the knowledge set is assigned as the parent patent. Thereby, a directed graph is generated linking each patent to its knowledge-based antecedent. Finally, with the help of this graph, the ratio of non-local to local patents as antecedents is investigated within focal sectors (and regions in the case of unrelated path development) to establish external and local knowledge sourcing. The next section elaborates further on the qualitative meso-level operationalising of knowledge.

3.3.2 Phase 2: Knowledge Networks and Knowledge Simulation

3.3.2.1 Knowledge Operationalization

As previously described, most studies focus on analysing knowledge in simplified vector and scalar quantities. Only a few studies represent knowledge elements as network structures (Cassi & Zirulia, 2008; Cowan & Jonard, 2004; Morone & Taylor, 2004; Schlaile et al., 2018). The study employs a network-of-networks approach, that means each agent (e.g., firm or individual) in an innovation network itself contains a network of knowledge units or skills (Schlaile et al., 2018). Knowledge is defined at two levels. While one is at the level of an agent (here, an organization), the other is at an aggregate regional level. In line with Morone and Taylor (2010), and Vermeulen and Pyka (2018, 2014b, 2014a), technological knowledge of an agent "is perceived as 'units' that are ordered in directed graphs, from primitive (upstream) to advanced (downstream)" (Vermeulen & Pyka, 2018,

p. 781). Patent data is extracted from the PATSTAT database which contains patent (application) data that can be further linked to geographical locations based on the addresses of inventors and applicants. Patents, as a proxy, are regarded as the most elementary (discrete) building blocks of an actor's knowledge-base in many studies that reconstruct an actor's knowledge-base out of patents (Buchmann, 2015; Fleming & Sorenson, 2001; Jaffe, 1989). The International Patent Classification (IPC) code is used to define the different units or skills. Elements of a knowledge-base are generally not independent from each other and there is relatedness between them. The knowledge network for each of the actors is reconstructed from patents filed by the actor, in moving time windows each encompassing five years.

Figure: Knowledge-base of a firm (schematic representation)

Co-occurrence of IPC subclasses help in figuring out the relatedness between skills as well as between the various industries and actors. This is used to reconstruct the knowledge base at a regional level, which in this study is referred to as the *regional skill profile*. The IPC sub-classes (4-digit level) are taken as nodes and ties are added between nodes whenever the subclasses co-occur on a patent (Buchmann, 2015; Saviotti, 2009). While studies measure aggregate knowledge through the mean average knowledge stocks of all agents (Cassi & Zirulia, 2008; Morone et al., 2007; Morone & Taylor, 2004; Zhuang et al., 2011), the qualitative aspects of skills, the structure of the knowledge-base and the and the methods of analysis for studying such structures have largely remained unexplored (Yayavaram & Ahuja, 2008). Therefore, this study plans to use network analysis to investigate the structure of the regional skill profile (Saviotti, 2009) including network measures such as centrality and modularity. In a modular knowledge-base, skills exist as clusters which are knit together (Buchmann, 2015). The modularity measure is expected to capture the interdependencies of knowledge units in innovation as well as in the emergence and survival of skills associated with the renewable sector. It is also expected that the current shift towards renewable energy becomes visible in the centrality measures of the skills associated with renewables in the regional skill profile. Finally, the study also employs the structural diversity measure proposed by Tom Broekel (2019) to define the complexity of a technology or knowledge base. This measure models technologies as combinatorial networks and derives a measure of technological complexity by measuring structural diversity of (sub-)network topologies in these networks. It allows for differentiating the different knowledge-bases (both at agent and regional levels) qualitatively according to the degree of complexity. Additionally, the stylized fact that the complexity of technology increases over time also hints at an increased role of related variety knowledge (given the high search costs associated with the unrelated variety knowledge for technologies with higher specialization) in the development of an industry over the years, which should not be conflated with an increased role of related knowledge in the late life-cycle of an industry. Therefore, this study controls for complexity of technology while investigating the drivers for network formation along the life cycle of an industry.

First, the knowledge base is reconstructed as a network. The IPC sub-classes (4-digit level) are taken as nodes and ties are added between nodes whenever the subclasses co-occur on a patent (Buchmann, 2015; Saviotti, 2009). This is done at an aggregate European level as well as at the level of each region (NUTS-3). Thereafter, an ego-centric network is delineated based on the sector in focus. The IPC Green Inventory is used to define the IPC codes to be considered as the central nodes for the ego-centric network. For example, the central nodes for the solar energy sector include– F24S, H02S, H01L etc. A critical threshold defined for relatedness is used to prune the network beyond a value to delineate the network for study. This network is also referred to as the Regional Skill Profile in this study. The Regional Skill Profile is dynamic and changes as new patents are granted. The underlying network structures is studied using social network analysis (SNA). To start with, modularity clusters are identified at the European scale. Thereafter, the shifting centrality of these clusters within the Regional Skill Profile is analysed over time and its effect on spatial diffusion and innovation dynamics is studied (the variables and their operationalization is mentioned in Table 3.2). Just as before, a Bayesian survival model is used to identify the determinants influencing the emergence of a sector or a patent within a region. Inter-dependencies between skills, in terms of their sustenance and their catalysing of spatial diffusion dynamics as well as innovation is also studied. This is analysed through a regression tree model. Moreover, the effect of the complexity of skills (in aggregate and within different modularity clusters) present in a region is analysed. Finally, the patents are distinguished between new or existing combinations based on whether a patent employs a new combination of classification codes or an existing one. The former is taken as a definition for radical novelty. A Stochastic Actor-Oriented Models (described further in the next section) is used to study the evolution of the Regional Skill Profile as well as the splitting and merging of skills.

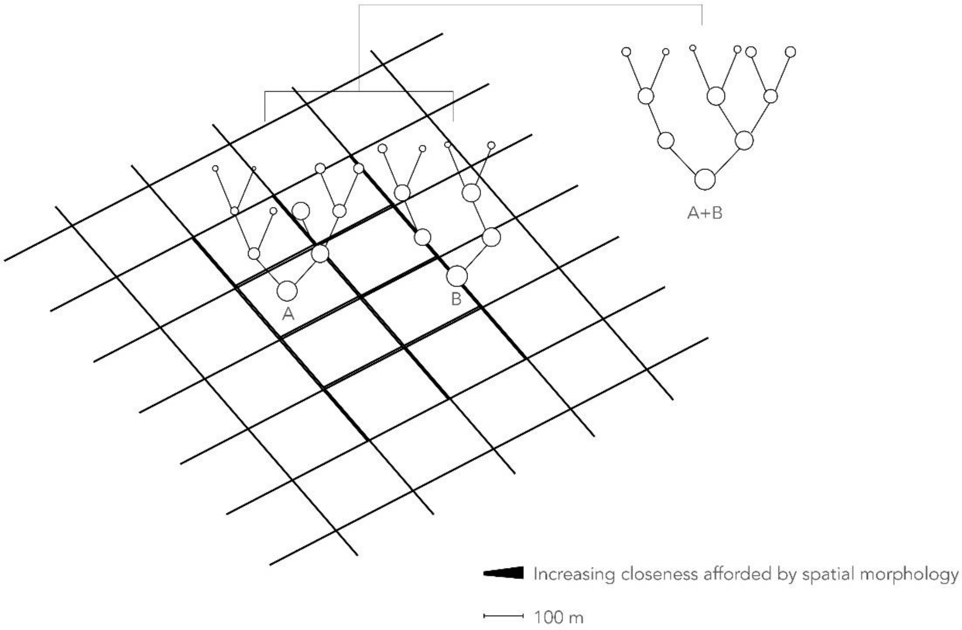


Figure: Regional Skill Profile is defined as the aggregate regional skill structure which is a product of individual knowledge-bases and interactions within them (schematic representation)

The knowledge network for each of the actors is also reconstructed from patents in moving time windows each encompassing five years. This study operationalizes concepts along the lines of Tobias Buchmann (2015). Technological distance w between two actors is calculated as:

$$w_{techdis\,ij} = \sqrt{\sum_{c=1}^{N} (p_i^c - p_j^c)^2}$$

where p_i^c and p_j^c are the patents granted to actors i and j respectively in patent category c. Further, in order to calculate modularity, the clustering coefficient cc_i of node (IPC sub-class) i with k_i links within the knowledge-base network is defined as below:

$$cc_i = \frac{n_i}{\frac{k_i \cdot (k_i - 1)}{2}}$$

where n_i is the number of ties between the k_i neighbors of node i. The denominator stands for the maximum number of ties which are possible amongst the k_i neighbors of node i.

Finally, the values of structural complexity given in Broekel (2019) are used to measure technological complexity. These complexity values are annual for each of the 645 four-digit CPC (Cooperative Patent Classification) classes (Juh□sz et al., 2020). Additionally, to calculate the complexity of the aggregate regional knowledge-base, the network diversity measure *iNDS* proposed by Broekel (Broekel, 2019) is used. The measure combines multiple network characteristics into one: It considers the share of modules α_{module} (densely connected subgraphs in a network), the variability of module sizes v_{module}, variability of the Laplacian matrix v_{λ} and relation of graphlets of size three and four $r_{graphlet}$.

$$iNDS(G_T) = \frac{\alpha_{module} * r_{graphlet}}{v_{module} * v_{\lambda}}.$$

Table 3.2: Knowledge related variables

Variable	Operationalization	Source
Low Technological Complexity	Number of skills with low technological complexity in region i (NUTS-3) at time t-3	Technological Complexity Measure based on four-digit CPC (Cooperative Patent Classification) classes, obtained from database made available by Broekel
Medium Technological Complexity	Number of skills with medium technological complexity in region i (NUTS-3) at time t-3	
High Technological Complexity	Number of skills with high technological complexity in region i (NUTS-3) at time t-3	
Modularity Cluster j	Number of skills in the modularity cluster j of the focal clean-energy sector at time t-3	Modularity clusters obtained from co-occurrence matrix in the PATSTAT database
Centrality-Modularity j	Average centrality measure for modularity cluster j within the regional skill profile at time t-3	Social Network Analysis to obtain centrality measure
Complexity-Modularity j	Average Technological Complexity measure for modularity cluster j within the regional skill profile at time t-3	PATSTAT
Patents	Number of Patents granted in region i (NUTS-3) at time t	PATSTAT

High Impact Patents	Number of top 10 percent most cited patents in region i (NUTS-3) at time t	PATSTAT
Non-Local Links - Modularity Cluster j	Percentage of non-local to local links for modularity cluster j	PATSTAT
New Combinations	Count of the number of new combinations of classification codes in region i at time t	PATSTAT

3.3.2.2 Knowledge Networks

In this study, we analyse the formation and evolution of knowledge networks along the life cycle of an industry as well as the effect of network measures on regional innovation. The latter is studied through investigating how network structure leads to new or existing combinations of knowledge elements.

In line with Balland et al. (2013) a stochastic actor-oriented model (Snijders, 2001) is employed to analyse the evolution of the collaboration network as well as the splitting and merging of skills. As Balland et. al. highlight, this approach allows for the simultaneous evaluation of three different driving forces: (1) individual characteristics, for example, the capacity to absorb knowledge; (2) relational structures; and (3) similarity between attributes of firms (like being proximate in cognitive or geographical terms). The empirical investigation of network dynamics require specific statistical models due to the presence of complex relational structures (Snijders, 2001). A fundamental property of network structures is the existence of conditional dependencies between observations. Balland et. al. (2013, p. 752) further note that "such network dependencies violate standard statistical procedures like OLS and logistic regressions that assume independence among observations." Therefore, a Markov chain Monte Carlo simulation model is considered fit to model network dependencies. Stochastic Actor-Oriented Models (SAOM) form the most promising class of models that allow for statistical inference of network dynamics (Snijders et al., 2008).

Along with this, questionnaires and personal interviews are conducted with clean energy firms in Netherlands to gather information on R&D activities of firms, their innovative success, and their engagement in knowledge exchange activities. The questionnaire is prepared on the lines of *'Constructing regional advantage: Towards state-of-the-art regional innovation policies in Europe?'* European project, in which research teams in a number of countries were interviewed regarding their knowledge exchange and innovation activities (Broekel & Boschma, 2017). These interviews were conducted by the word of mouth and the questions were adapted for the corresponding industry. However, this database did not deal with the renewable sector. On similar lines, detailed information is sought to be collected on all kinds of organizations (other firms, associations, universities, etc.) that a firm interacted with in the last three years and with whom they shared knowledge relevant for their innovation activities. To augment the dynamic analysis elucidated above, a static analysis is conducted in order to explain the formation of the present network structure. Along the lines of Broekel and Boschma (2017), a cluster analysis technique will be employed to study the interaction effects between the cognitive and geographical proximity dimension by systematically identifying groups of firms with specific geographical and cognitive characteristics. These groups are then tested separately to see which ones are more conducive for innovative performance.

Moreover, other information is sought including the various skills involved in knowledge exchange to establish the underlying mechanism of knowledge diffusion. It is assumed that different skills within the focal sector are likely to have a varying degree of diffusion. Further, different actors may adopt different collaboration strategies, thereby impacting network formation dynamics. This information is also used to calibrate the agent based simulation model elaborated in the next section.

3.3.2.3 Knowledge Simulation

An agent-based simulation model is used to simulate aggregate and individual level knowledge production and in turn the innovation propensity of a region. This model is validated over the data obtained for the previous five years, from not only sources such as PATSTAT, that capture formal collaborations and patentable innovations, but also personal interviews and questionnaires that capture informal interactions as well as technologies that are not always patentable. Thereafter, it used to simulate knowledge production and innovation within various regions in the near future.

As described previously, the knowledge network for each of the actors is reconstructed from patents in moving time windows each encompassing five years. The knowledge base of each agent is arranged in 'tiers', the equivalent of years in the patent database. After the knowledge-base of each actor is defined at time t as a directed graph, the top three tiers of the graph are considered *active* for knowledge exchange. Local collaboration network between agents is deduced from collaborations on patents and questionnaires (or alternately through Dutch subsidies catalogue or the CORDIS database) and set as the collaboration network at time t. Thereafter, the agents in a network are allowed to exchange knowledge as well as enter into a collaboration or exit from one, based on certain stylized facts as obtained from the previous analyses and interviews. In line with Morone and Taylor (2010), and Vermeulen and Pyka (2018, 2014b, 2014a), "…one agent or a collective of agents tries to unlock units in this directed graph by combining units that it already possesses" (Vermeulen & Pyka, 2018, p. 781). Every agent is assigned a visibility parameter of possible skills that can be unlocked based on its knowledge-base and its structural position in the network, that establishes the probability of new skills emerging (through merging or splitting) in the environment. Based on the relevance of geographical, cognitive and social proximity, as obtained from previous analyses, the neighbourhood of an agent is defined in these three dimensions, with a higher probability of collaboration and knowledge exchange in the neighbourhood. This is aimed to capture whether agents in a system lock themselves into a state of local but not global maxima, thereby allowing us to understand if (new) knowledge creation dynamics are being fully exploited in a system or not.

At every time step new knowledge in terms of new skills is generated at both the individual as well as the aggregate level, and certain skills are rendered obsolete. New skills are qualitatively defined in terms of their technological complexity, modularity, centrality and radicalness and Regional Skill Profile is generated. Thereby, inferences are made on the spatial emergence and concentration of the clean-energy sector and the regional innovation.

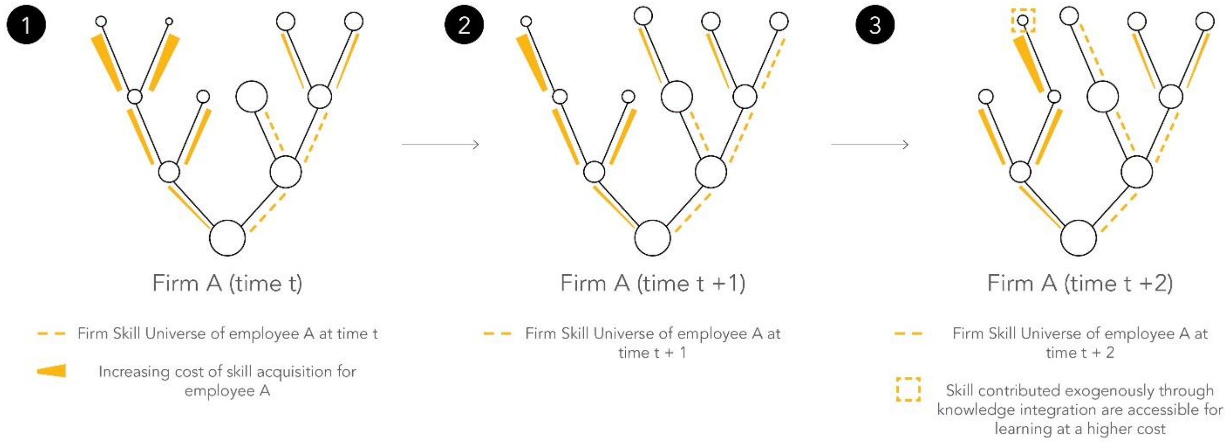


Figure: Skill Universe of a single agent and its change with time (schematic representation)

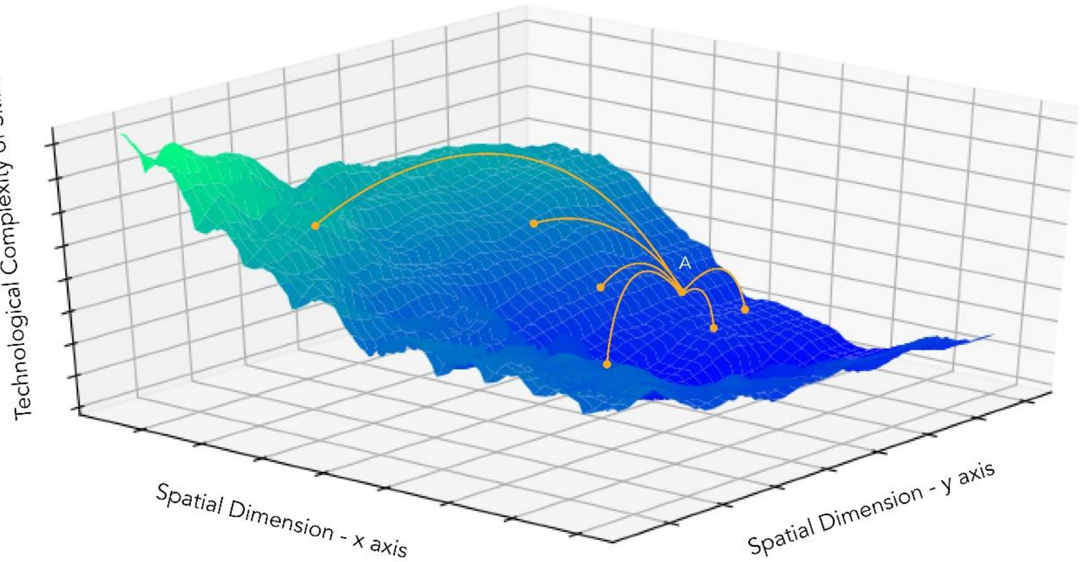


Figure: Aggregate regional knowledge in space, agent A's neighbourhood is defined in space as well as through its social network (here, depicted as yellow links), we also observe how certain areas in the graph are locked into a state of local maxima (schematic representation)

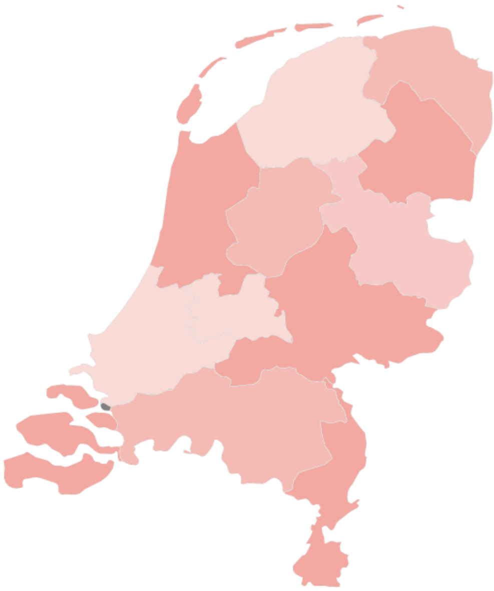


Figure: Relative potential of (new) knowledge production as expected outcome (schematic representation)

3.3.3 Phase 3: Resilience

Using the agent based simulation model, as described in section 3.3.2.3, the effect of firm mortality and entry (and the corresponding skill removal and introduction) on the regional innovation system is analysed. Agents and corresponding skills (nodes) are removed and introduced systematically to the regional knowledge network and the effects on the robustness, efficiency of the knowledge network and aggregate knowledge production is analysed. It is likely that the push towards localization of supply change might non-linearly affect certain regional innovation systems given not only how the new skills might add to the heterogeneity of the system, but also how they contribute to the overall modularity and technological complexity of the system. Further, considering that our previous analysis disentangles the relevance of local and related knowledge in the development of an industry (supply-side), possible scenarios are generated for the varying possibility of development and decline of industries in different regions under the influence of an external (economic) shock, taking business-as-usual as the benchmark scenario. The data on firm mortality and supply-chain localization is gathered through a detailed web research. Moreover, the impact of the current pandemic on collaborations and innovation performance is assessed through questionnaires and personal interviews, thereby, informing the simulation model.

Finally, the study looks beyond the focal sectors and investigates the cascading effects of firm mortality and entry on the regional knowledge networks and regional innovation systems of the related industries. The impact of the current pandemic shock on inter-sectoral collaborations is assessed through the above mentioned questionnaire. The resulting regional skill profile is used to simulate the impact of the external (economic) shock on the resilience of the region.

3.4 Overview of Research Questions and Methods

The following table summarizes the methodology and expected output for each of the research question.

<table>
<tr><td colspan="3">Research question 1: What are the key factors responsible for the spatial emergence & concentration of the clean energy sector?</td></tr>
<tr><td>Guiding Sub-questions</td><td>Method and Data</td><td>Expected Output</td></tr>
<tr><td>What is the role of local and non-local knowledge across technological fields in the emergence and evolution of the various clean-energy sectors?</td><td rowspan="3">Method: OLS regression with explanatory variables lagged in time/ Bayesian spatial survival analysis

Data: PATSTAT/ OECD REGPAT/ Bureau van Dijk database for firm location and knowledge related variables; OSM and web research for estimating the number of solar panels, wind turbines, EV chargers within a region; Manifesto Project Dataset to proxy policy support for renewables</td><td>It is believed that both the local knowledge production and the non-local knowledge sourcing are crucial for the emergence and concentration of industries; a higher technological complexity is critical for solar and EV industries</td></tr>
<tr><td>What is the industry's (specifically local knowledge's) role in developing local demand (supply-push)?</td><td>Geographical proximity to manufacturers and high local knowledge production are believed to have a positive impact on the development of local demand</td></tr>
<tr><td>What is the relevance of local demand (in conjunction with other clean energy sectors and constraining factors) and supply in the geographic clean-energy sector niche and regime formation?</td><td>While it is speculated that presence of other renewable energy sectors within a region may have a positive impact on the emergence of clean energy sector; the presence of fossil fuel industry in the region might act as a deterrent</td></tr>
<tr><td colspan="3">Research question 2: What are the key driving factors behind the formation & evolution of knowledge networks within the clean energy sector?</td></tr>
<tr><td>Guiding Sub-questions</td><td>Method and Data</td><td>Expected Output</td></tr>
</table>

<table>
<tr>
<td>How does the quality of skills (such as their centrality, modularity and technological complexity) affect regional innovation?</td>
<td rowspan="3">Defining modularity clusters both at a European and a regional scale using PATSTAT database, defining regional skill network based on co-occurrences on patents; assigning average centrality and technological complexity measure to each modularity cluster followed by OLS regression with patents generated and radical novelty; questionnaires and personal interviews to determine the knowledge network, innovation output of a collaboration, mechanism of knowledge diffusion; Stochastic Actor-Oriented Models (SAOM) using SIENA statistical package on R to determine the driving factors behind network evolution; cluster analysis to group firms based on their cognitive and geographical proximity followed by OLS regression; agent-based simulation using NetLogo to simulate aggregate knowledge generation, calculation of technological complexity, modularity and novelty of the resulting aggregate regional knowledge
Data: PATSTAT, Questionnaires and Interviews, Technological Complexity Measure</td>
<td>Certain modularity clusters are considered to be more conducive to catalysing innovation, however, their structural position in the regional skill network is expected to matter</td>
</tr>
<tr>
<td>What is the role of network measures such as proximity, modularity, embeddedness and status, local and non-local knowledge across technological fields in the formation of knowledge links and subsequent knowledge exchange between actors at various scales?</td>
<td>The geographical and cognitive proximity are believed to act as both substitutive and complementary based on the spatial scale of observation; local knowledge is speculated to impact the evolution of knowledge network, a higher technological complexity is more conducive to a denser network</td>
</tr>
<tr>
<td>How to simulate aggregate regional knowledge production?</td>
<td>Such a simulation could help in predicting the innovation propensity of the renewable sector in a region as well as the probability of the region to diversify or sustain the renewable sector</td>
</tr>
<tr>
<td colspan="3">Research question 3: How do external (economic) shocks affect the knowledge production and spatial diffusion dynamics of the clean energy sector, and what role does the sector play in regional resilience?</td>
</tr>
<tr>
<td>Guiding Sub-questions</td>
<td>Method</td>
<td>Expected Output</td>
</tr>
<tr>
<td>How does firm mortality/ entry affect aggregate regional knowledge production within the various clean-energy sectors?</td>
<td>Agent-based simulation to estimate the impact on knowledge generation;</td>
<td>Relative difference between the various regions in their knowledge production capacity when exposed to external (economic) shock</td>
</tr>
</table>

How does firm mortality/ entry affect the geographical configurations of niche formation and regimes within the various clean-energy sectors?	OLS regression to estimate the impact on spatial diffusion processes	Relative difference between the various regions in their capacity to diversify or sustain clean energy sectors when exposed to external (economic) shock
What are the cascading effects of firm mortality/ entry within the clean energy sector on related industries?		Certain sectors would be impacted more than others, and the dependencies between modularity clusters is expected to indicate this

3.5 Data

The following table summarizes the datasets that the study employs.

Name	Description	Granularity
PATSTAT	PATSTAT contains bibliographical and legal event patent data from leading industrialized and developing countries	The data is available at a time granularity of one year and the spatial granularity of NUTS 3 regions
AMADEUS - Bureau Van Dijk	Primary source for private companies related data, contains micro-data on firms for all European countries, ownership information 4-digit NACE codes, size information in the form of number of employees and turnover	The data is available at a time granularity of one year and the spatial granularity of NUTS 3 regions, however location of the private companies can be georeferenced
CORDIS dataset	This dataset contains projects and related organizations funded by the European Union under the FP7 and Horizon 2020 framework program for research and innovation from 2008 to 2020	The data is available at a time granularity of one year and the spatial granularity of NUTS 2 regions
Dutch Provincial Data on Clean-Tech	Regional data on the clean energy sector within Netherlands	Varying time granularity and the spatial granularity of NUTS 2 regions
Technological Complexity Measure	These complexity values are annual for each of the 645 four-digit CPC (Cooperative Patent Classification) classes	The data is available annually

Eurostat	Europe-wide statistics and indicators	The data is available at a time granularity of one year and the spatial granularity of NUTS 2 regions
APIs – OSM, Google, Four Square	POI database, can be used for geocoding	Present day geocoded data
Questionnaires	Questionnaires to determine the informal knowledge networks, relatedness and innovation output of a collaboration	Actor's micro-spatial location specified
Personal Interviews	Personal interviews to establish the underlying mechanism of knowledge diffusion (by investigating the various skills involved in knowledge exchange)	Actor's micro-spatial location specified
Others	Dutch subsidies catalogue, Manifesto Project Database	-

3.6 Research Ethics

In the table below, I illustrate the ways each dataset will be stored and protected. I also list the associated ethical concerns that the mishandling of each dataset may yield. Considering that a major part of the study intends to make use of secondary data types (including PATSTAT, CORDIS, Bureau Van Dijk etc.), the analytical rigour and accuracy, for the large part, can be easily verified. Where primary data is intended to be collected, the research protocol is described to ensure transparency. The management of this dataset is detailed below.

Primary data will be collected through questionnaires and personal interviews. The project intends to collect organisation-level data and no personal data will be collected. Data will be gathered from the volunteer participants (organizations) through a web app or in person, and they will be provided with an informed consent form, written in an easily understandable language. The form will inform the participants that their participation is voluntary and they have the right to withdraw at any time. The form will also contain a short description of the purpose and procedure of the research, and how it handles the collected data, in terms of storage and protection. The data will be stored in compliance with EU legislation and the data management policy in place at the University of Twente, and partner institutions if applicable. Data aggregation will be ensured for dissemination.

Finally, I line with the Netherlands (VSNU) guidelines, the data will be stored for a minimum of 10 years to ensure the verifiability of the research output. Considering the confidentiality of the data, only me and my supervisors would have access to it during the research process. However, access to anonymised meta-data and research protocols could be made available to third parties. The ethics questionnaire, further detailing these points, is intended to be filled after the qualifier.

Table: Data Management Plan

Datasets	Data Type // Classification	Personal Data	Sensitivity	Access	Ethical Concerns	Post-research protection and access
Questionnaires	Primary // Confidential	Yes	High	Researcher & supervisors	Privacy, anonymity, consent, bias, beneficence, misinterpretation	Restricted access & anonymised
Transcribed Interviews	Primary // Confidential	Yes	High	Researcher & supervisors	Privacy, anonymity, consent, bias, beneficence, misinterpretation	Restricted access & anonymised
Audio Recordings	Primary // Confidential	Yes	High	Researcher & supervisors	Privacy, anonymity, consent, bias, beneficence, misinterpretation	Restricted access & anonymised
Dutch Provincial Data on Clean-Tech	Secondary // Confidential	Yes	High	Researcher & supervisors	Privacy, anonymity, consent, bias, beneficence, misinterpretation	Restricted access & anonymised
PATSTAT	Secondary // Public	Yes	Medium	Researcher & supervisors	Bias, beneficence, misinterpretation	-
AMADEUS - Bureau Van Dijk	Secondary // Confidential	Yes	Medium	Researcher & supervisors	Bias, beneficence, misinterpretation	-
CORDIS dataset	Secondary // Public	Yes	Medium	Researcher & supervisors	Bias, beneficence, misinterpretation	-
Eurostat	Secondary // Public	No	Low/ Medium	Researcher & supervisors	Bias, beneficence, misinterpretation	-
APIs – OSM, Google, Four Square	Secondary // Public	No	Low/ Medium	Researcher & supervisors	Privacy, anonymity, bias, beneficence, misinterpretation	-
Technological Complexity Measure	Secondary // Public	No	Low	Researcher & supervisors	Beneficence, misinterpretation	-

4. Contribution of the Research

<table>
<tr><th>Research Gaps</th><th>Contribution of the Study</th></tr>
<tr><td>1. Role of multi-scalar knowledge dynamics and local demand in the emergence and evolution of the various clean energy sectors
Highlighted by: Hassink et al. (Bednarz & Broekel, 2020; Esposito, 2020; Hassink et al., 2019)</td><td rowspan="2">With the example of the clean energy sector, the study uses Bayesian survival analysis to empirically analyse the role of these factors, it uses the unique dataset prepared by Christopher Esposito (Esposito, 2020) to operationalize external knowledge sourcing</td></tr>
<tr><td>2. Co-evolutionary dynamics between local demand and local knowledge production in the development of the various clean-energy sectors
Highlighted by: Bednarz and Broekel (2020)</td></tr>
<tr><td>3. Role of network measures such as modularity, centrality and technological complexity of knowledge elements in the evolution of the various clean energy sectors and regional innovation
Highlighted by: Buchmann, Bednarz and Broekel (Bednarz & Broekel, 2020; Buchmann, 2015)</td><td>Considering the importance of the mentioned network measures, the study uses network analysis to qualitatively describe aggregate knowledge in these terms and subsequently uses OLS regression to investigate the interdependencies of skills as well as their catalysing of innovation</td></tr>
<tr><td>4. Empirical research on knowledge network evolution within the various clean energy sectors, and the role of geography in it
Highlighted by: Tsouri et al. (Balland et al., 2013; Tsouri et al., 2020)</td><td>The study uses Stochastic Actor Oriented Model (SAOM) to study the evolution of knowledge networks within the clean energy sector while keeping the spatial dimension and above mentioned network measures at the centre of the analysis</td></tr>
<tr><td>5. Agent based simulation to arrive at individual and aggregate regional knowledge indicating the propensity for innovation for a certain region
Highlighted by: Pyka et. al. (Buchmann, 2015; Pyka et al., 2009; Schlaile et al., 2018; Vermeulen & Pyka, 2017)</td><td>The study makes advancements to knowledge operationalisation through the incorporation of network measures as well as empirically calibrated innovation reward landscape in the system based on the how combinations of skills (new vs existing) affect regional innovation</td></tr>
</table>

6. Resilience of the knowledge production within the various clean-energy sectors when exposed to an (external) economic shock and subsequent impact on the niche and regime formation dynamics Highlighted by: Tsouri et al. (Toth et al., 2020; Tsouri & Pegoretti, 2020)	Through the previous ABM model, the study analyses the introduction and removal of knowledge elements (based on the entry and exit of firms due to an exogenous shock) from the system and investigates the change in aggregate knowledge production and spatial diffusion dynamics, and in turn the resilience of a region

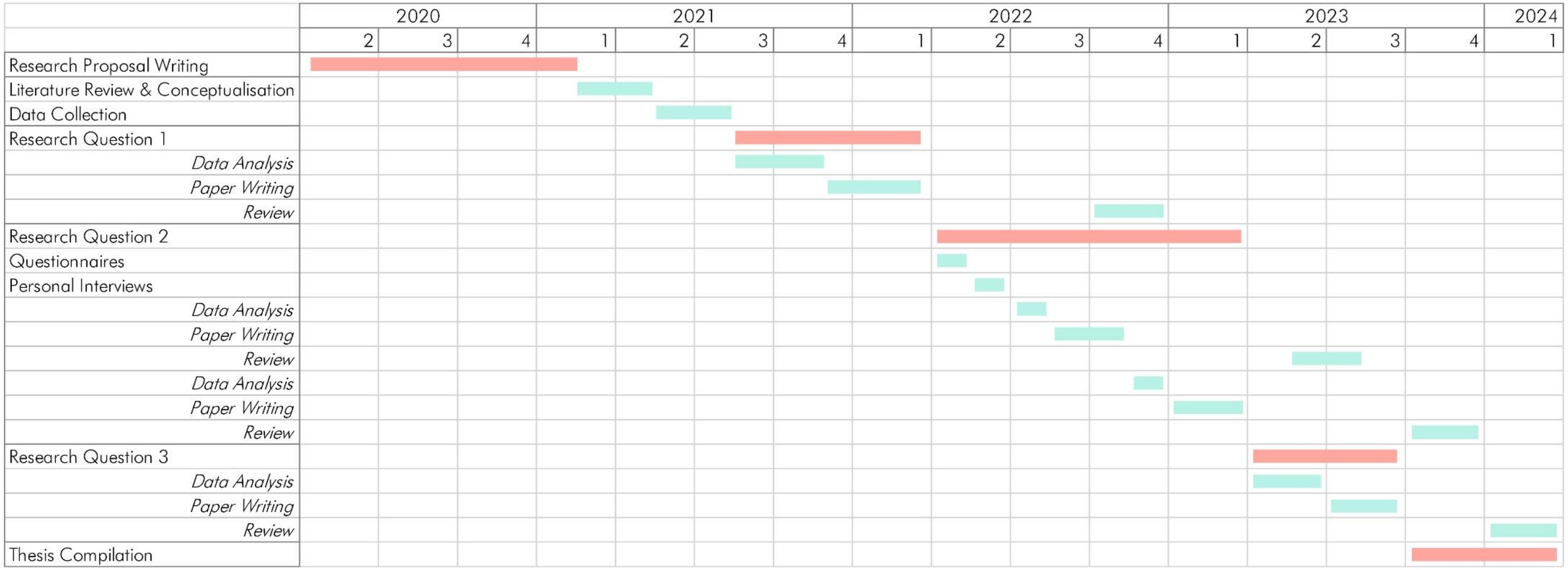
5. Research Schedule

5.1 Research Timeline

The next steps in the research are twofold. First, the theoretical and conceptual approach has to be developed further. While the proposal broadly specifies the approach, further work needs to be done in order to apply the conceptual framework to the specific case. Moreover, while the proposal details the knowledge operationalization at various scales, there has to be further calibration based on the insights gained from the personal interactions with the industry. Second, the empirical data needs to be collected. Certain datasets can be accessed online whereas others are required to be physically collected (personal interviews). However, as described in the proposal, given that the first phase of the study involves datasets which can be accessed online, the work can start immediately. Personal interviews are planned as a part of the second phase of the study, given the ongoing Covid-19 pandemic. If there is a reasonable improvement in the situation, the interviews could be conducted face-to-face, as initially planned, or otherwise they can be substituted with online interviews to avoid any significant delays.

As is outlined previously, research question 1 and 3 correspond to a focus for a journal article each. Research question 2, on the other hand, is expected to be the focus for two journal articles. The research timeline, as outlined in the table below, demarcates time for secondary data collection, data analysis (roughly 1 quarter for each article), paper writing (roughly 1 quarter for each article), questionnaires, personal interviews and finally thesis compilation. Additional training would also be acquired during the course of the PhD which would include discipline subjects (multi agent systems, non-linear dynamics, spatio-temporal analytics and modeling etc.) and generic subjects (data management bootcamp, academic publishing bootcamp, academic presentations bootcamp, scientific information bootcamp etc.) as is outlined in the table below.

Table: 5.1 Research Timeline


2020
2021
2022
2023
2024
2 3 4 1 2 3 4 1 2 3 4 1 2 3 4 1
Research Proposal Writing
Literature Review & Conceptualisation
Data Collection
Research Question 1
Data Analysis
Paper Writing
Review
Research Question 2
Questionnaires
Personal Interviews
Data Analysis
Paper Writing
Review
Data Analysis
Paper Writing
Review
Research Question 3
Data Analysis
Paper Writing
Review
Thesis Compilation

5.2 Training and Supervision Plan

Table 5.2: Training and Supervision Plan

Activity	ECTS	Completed
Generic Subjects		
Mandatory first year generic PhD Courses		
TGS Introductory Workshop + Academic Integrity	1.5	1.5
Academic Publishing	2.0	
Academic Presentations	1.5	
Data Management	1.0	
Scientific Information	0.5	
Research Support		
Science writing	2.0	
How to improve your scientific footprint	0.5	
Competing for grants	0.5	
Language Course		
Various Dutch courses	Max 5	
Discipline Subjects		
Multi Agent Systems	5	
Non-Linear Dynamics: Mathematical and Computational Approaches	2	
Spatio-Temporal Analytics and Modeling	7	
Statistics for Spatial and Spatio-Temporal Data	7	
Conferences and Workshops	2	
Total	33.5	

6. References

6.1 List of References

Allen, T. J. (1977). Managing the flow of technology: Technology transfer and the dissemination of technological information within the R & D organization(Book). *Research Supported by the National Science Foundation. Cambridge, Mass., MIT Press, 1977. 329 P*.

Asheim, B. T., Boschma, R., & Cooke, P. (2011). Constructing regional advantage: Platform policies based on related variety and differentiated knowledge bases. *Regional Studies*, *45*(7), 893–904.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American Economic Review*, *86*(3), 630–640.

Axelrod, R. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration* (Vol. 3). Princeton university press.

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, *65*(3), 595–621.

Balland, P.-A. (2012). Proximity and the evolution of collaboration networks: evidence from research and development projects within the global navigation satellite system (GNSS) industry. *Regional Studies*, *46*(6), 741–756.

Balland, P.-A., Belso-Mart'nez, J. A., & Morrison, A. (2016). The dynamics of technical and business knowledge networks in industrial clusters: Embeddedness, status, or proximity? *Economic Geography*, *92*(1), 35–60.

Balland, P.-A., Boschma, R., & Ravet, J. (2019). Network dynamics in collaborative research in the EU, 2003–2017. *European Planning Studies*, *27*(9), 1811–1837.

Balland, P.-A., De Vaan, M., & Boschma, R. (2013). The dynamics of interfirm networks along the industry life cycle: The case of the global video game industry, 1987–2007. *Journal of Economic Geography*, *13*(5), 741–765.

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., & Hidalgo, C. A. (2020). Complex economic activities concentrate in large cities. *Nature Human Behaviour*, *4*(3), 248–254.

Bednarz, M., & Broekel, T. (2020). Pulled or pushed? The spatial diffusion of wind energy between local demand and supply. *Industrial and Corporate Change*.

Binz, C., & Truffer, B. (2017). Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts. *Research Policy*, *46*(7), 1284–1298.

Boschma, R. (2005). Proximity and innovation: a critical assessment. *Regional Studies*, *39*(1), 61–74.

Boschma, R. (2015). Towards an evolutionary perspective on regional resilience. *Regional Studies*, *49*(5), 733–751.

Boschma, R. (2017). Relatedness as driver of regional diversification: A research agenda.

Regional Studies, *51*(3), 351–364.

Boschma, R., Coenen, L., Frenken, K., & Truffer, B. (2017). Towards a theory of regional diversification: combining insights from Evolutionary Economic Geography and Transition Studies. *Regional Studies*, *51*(1), 31–45.

Boschma, R., Frenken, K., Bathelt, H., Feldman, M., & Kogler, D. (2012). Technological relatedness and regional branching. *Beyond Territory. Dynamic Geographies of Knowledge Creation, Diffusion and Innovation*, 64–68.

Boschma, R., Minondo, A., & Navarro, M. (2013). The emergence of new industries at the regional level in s pain: A proximity approach based on product relatedness. *Economic Geography*, *89*(1), 29–51.

Boschma, R A. (n.d.). K.. Frenken. 2010. The Spatial Evolution of Innovation Networks: A Proximity Perspective. *The Handbook of Evolutionary Economic Geography. Cheltenham: Edward Elgar*.

Boschma, Ron A, & Frenken, K. (2006). Why is economic geography not an evolutionary science? Towards an evolutionary economic geography. *Journal of Economic Geography*, *6*(3), 273–302.

Boschma, Ron A, & Lambooy, J. G. (1999). Evolutionary economics and economic geography. *Journal of Evolutionary Economics*, *9*(4), 411–429.

Boschma, Ron A, & Wenting, R. (2007). The spatial evolution of the British automobile industry: Does location matter? *Industrial and Corporate Change*, *16*(2), 213–238.

Breul, M., Broekel, T., & Brachert, M. (2015). The drivers of the spatial emergence and clustering of the photovoltaic industry in Germany. *Zeitschrift F□r Wirtschaftsgeographie*, *59*(3), 133–150.

Broekel, T. (2019). Using structural diversity to measure the complexity of technologies. *PloS One*, *14*(5), e0216856.

Broekel, T., & Boschma, R. (2012). Knowledge networks in the Dutch aviation industry: the proximity paradox. *Journal of Economic Geography*, *12*(2), 409–433.

Broekel, T., & Boschma, R. (2017). *The cognitive and geographical structure of knowledge links and how they influence firms' innovation performance*.

Buchmann, T. (2015). *The evolution of innovation networks: an automotive case study*. Springer.

Burton, T., Sharpe, D., Jenkins, N., & Bossanyi, E. (2001). *Wind energy handbook* (Vol. 2). Wiley Online Library.

Cantner, U., Graf, H., & Hinzmann, S. (2015). *The role of geographical proximity for project performance: Evidence from the German" Leading-Edge Cluster Competition"*. Jena Economic Research Papers.

Carayol, N., & Roux, P. (2009). Knowledge flows and the geography of networks: A strategic model of small world formation. *Journal of Economic Behavior & Organization*, *71*(2), 414–427.

Cassi, L., & Plunket, A. (2010). *The determinants of co-inventor tie formation: proximity and network dynamics*.

Cassi, L., & Plunket, A. (2015). Research collaboration in co-inventor networks: combining closure, bridging and proximities. *Regional Studies*, *49*(6), 936–954.

Cassi, L., & Zirulia, L. (2008). The opportunity cost of social relations: on the effectiveness of small worlds. *Journal of Evolutionary Economics*, *18*(1), 77–101.

Coenen, L., & Truffer, B. (2012). Places and spaces of sustainability transitions: geographical contributions to an emerging research and policy field. *European Planning Studies*, *20*(3), 367–374.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative Science Quarterly*, 128–152.

Cowan, R., & Foray, D. (1997). The economics of codification and the diffusion of knowledge. *Industrial and Corporate Change*, *6*(3), 595–622.

Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, *28*(8), 1557–1575.

Crespo, J., Suire, R., & Vicente, J. (2014). Lock-in or lock-out? How structural properties of knowledge networks affect regional resilience. *Journal of Economic Geography*, *14*(1), 199–219.

David, P. F. (2002). D.(2002)."And Introduction to the Economy of the Knowledge." *International Social Science Journal*, *171*.

Edquist, C., & Johnson, B. (1997). System of innovation: overview and basic concepts. *Edqvist. C.*

Ellison, G., & Fudenberg, D. (1993). Rules of thumb for social learning. *Journal of Political Economy*, *101*(4), 612–643.

Ellison, G., & Fudenberg, D. (1995). Word-of-mouth communication and social learning. *The Quarterly Journal of Economics*, *110*(1), 93–125.

Esposito, C. (2020). *The Emergence of Knowledge Production in New Places*. Utrecht University, Department of Human Geography and Spatial Planning ….

Fagiolo, G., Moneta, A., & Windrum, P. (2006). Confronting agent-based models with data: methodological issues and open problems. In *Advances in Artificial Economics* (pp. 255–267). Springer.

Feldman, M. P., & Florida, R. (1994). The geographic sources of innovation: technological infrastructure and product innovation in the United States. *Annals of the Association of American Geographers*, *84*(2), 210–229.

Fleming, L., & Sorenson, O. (2001). Technology as a complex adaptive system: evidence from patent data. *Research Policy*, *30*(7), 1019–1039.

Foray, D. (2014). From smart specialisation to smart specialisation policy. *European Journal of Innovation Management*.

Fornahl, D., Hassink, R., Klaerding, C., Mossig, I., & Schr□der, H. (2012). From the old path of shipbuilding onto the new path of offshore wind energy? The case of northern Germany. *European Planning Studies*, *20*(5), 835–855.

Frenken, K. (2006). Technological innovation and complexity theory. *Economics of Innovation and New Technology*, *15*(2), 137–155.

Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, *41*(5), 685–697.

Garud, R., & Karn□e, P. (2003). Bricolage versus breakthrough: distributed and embedded agency in technology entrepreneurship. *Research Policy*, *32*(2), 277–300.

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Research Policy*, *31*(8–9), 1257–1274.

Geels, F. W. (2004). From sectoral systems of innovation to socio-technical systems: Insights about dynamics and change from sociology and institutional theory. *Research Policy*, *33*(6–7), 897–920.

Gereffi, G. (2020). What does the COVID-19 pandemic teach us about global value chains? The case of medical supplies. *Journal of International Business Policy*, *3*(3), 287–301.

Geroski, P. A. (2000). Models of technology diffusion. *Research Policy*, *29*(4–5), 603–625.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or the undefinable tacitness of being (there). *Journal of Economic Geography*, *3*(1), 75–99.

Giuliani, E. (2013). Network dynamics in regional clusters: Evidence from Chile. *Research Policy*, *42*(8), 1406–1419.

Hassink, R., Isaksen, A., & Trippl, M. (2019). Towards a comprehensive understanding of new regional industrial path development. *Regional Studies*, *53*(11), 1636–1645.

Heiberg, J., Binz, C., & Truffer, B. (2020). *The Geography of Technology Legitimation. How multi-scalar legitimation processes matter for path creation in emerging industries*. Utrecht University, Department of Human Geography and Spatial Planning ….

Howells, J. R. L. (2002). Tacit knowledge, innovation and economic geography. *Urban Studies*, *39*(5–6), 871–884.

Jaffe, A. B. (1989). Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. *Research Policy*, *18*(2), 87–97.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577–598.

Juh□sz, S., Broekel, T., & Boschma, R. (2020). Explaining the dynamics of relatedness: The role of co-location and complexity. *Papers in Regional Science*.

Klepper, S. (2006). The evolution of geographic structure in new industries. *Revue de l'OFCE*, *5*, 135–158.

Klepper, S. (2002). The evolution of the US automobile industry and Detroit as its capital. *9th Congress of the International Joseph A. Schumpeter Society, Gainesville, Florida, March*, 22–23.

Kogler, D. F. (2015). *Evolutionary economic geography–Theoretical and empirical progress*. Taylor & Francis.

Kudic, M. (2014). *Innovation networks in the German laser industry: Evolutionary change, strategic positioning, and firm innovativeness*. Springer.

Malerba, F., & Torrisi, S. (1992). Internal capabilities and external networks in innovative activities. Evidence from the software industry. *Economics of Innovation and New Technology*, *2*(1), 49–71.

Martin, R., & Sunley, P. (2006). Path dependence and regional economic evolution. *Journal of Economic Geography*, *6*(4), 395–437.

Meder, A. (2008). *Technological and geographical patterns in the choice of cooperation partner*. Jena economic research papers.

Morgan, K. (2004). The exaggerated death of geography: learning, proximity and territorial innovation systems. *Journal of Economic Geography*, *4*(1), 3–21.

Morgan, K. (2015). *Smart specialisation: Opportunities and challenges for regional innovation policy*. Taylor & Francis.

Morone, A., Morone, P., & Taylor, R. (2007). A laboratory experiment of knowledge diffusion dynamics. In *Innovation, industrial dynamics and structural transformation* (pp. 283–302). Springer.

Morone, P., & Taylor, R. (2004). Knowledge diffusion dynamics and network properties of face-to-face interactions. *Journal of Evolutionary Economics*, *14*(3), 327–351.

Morone, P., & Taylor, R. (2010). *Knowledge diffusion and innovation: modelling complex entrepreneurial behaviours*. Edward Elgar Publishing.

Morris, S. (2000). Contagion. *The Review of Economic Studies*, *67*(1), 57–78.

Mueller, M., Bogner, K., Buchmann, T., & Kudic, M. (2017). The effect of structural disparities on knowledge diffusion in networks: an agent-based simulation model. *Journal of Economic Interaction and Coordination*, *12*(3), 613–634.

Neffke, F., Hartog, M., Boschma, R., & Henning, M. (2018). Agents of structural change: The role of firms and entrepreneurs in regional diversification. *Economic Geography*, *94*(1), 23–48.

Neffke, F., Henning, M., Boschma, R., Lundquist, K.-J., & Olander, L.-O. (2011). The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies*, *45*(1), 49–65.

Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., & Van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research Policy*, *36*(7), 1016–1034.

Podolny, J. M. (2001). Networks as the pipes and prisms of the market. *American Journal of Sociology*, *107*(1), 33–60.

Polanyi, M. (1967). The tacit dimension. Anchor. *Garden City, NY*.

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. *Papers in Regional Science*, *86*(3), 423–443.

Pyka, A., & Fagiolo, G. (2007). 29 Agent-based modelling: a methodology for neo-Schumpeterian economics'. *Elgar Companion to Neo-Schumpeterian Economics*, *467*.

Pyka, A., Gilbert, N., & Ahrweiler, P. (2009). Agent-based modelling of innovation networks–the fairytale of spillover. In *Innovation networks* (pp. 101–126). Springer.

Pyka, A., Kudic, M., & M□ller, M. (2019). Systemic interventions in regional innovation systems: entrepreneurship, knowledge accumulation and regional innovation. *Regional Studies*, *53*(9), 1321–1332.

Rigby, D. L. (2015). Technological relatedness and knowledge space: entry and exit of US cities from patent classes. *Regional Studies*, *49*(11), 1922–1937.

Roesler, C., & Broekel, T. (2017). The role of universities in a network of subsidized R&D collaboration: The case of the biotechnology-industry in Germany. *Review of Regional Research*, *37*(2), 135–160.

Saviotti, P P. (1999). Knowledge, information and organisational structures. In *Authority and Control in Modern Industry* (pp. 132–151). Routledge.

Saviotti, Pier Paolo. (1998). On the dynamics of appropriability, of tacit and of codified knowledge. *Research Policy*, *26*(7–8), 843–856.

Saviotti, Pier Paolo. (2009). Knowledge networks: structure and dynamics. In *Innovation Networks* (pp. 19–41). Springer.

Saxenian, A. (1994). *Regional networks: industrial adaptation in Silicon Valley and route 128*.

Schlaile, M. P., Zeman, J., & Mueller, M. (2018). It'sa match! Simulating compatibility-based learning in a network of networks. *Journal of Evolutionary Economics*, *28*(5), 1111–1150.

Schot, J., & Geels, F. W. (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technology Analysis & Strategic Management*, *20*(5), 537–554.

Scott, A. J., & Storper, M. (1986). *High technology industry and regional development: a theoretical critique and reconstruction*. Department of Geography, University of Reading.

Singh, J. (2005). Collaborative networks as determinants of knowledge diffusion patterns. *Management Science*, *51*(5), 756–770.

Slicher van Bath, B. H. (1963). *agrarian history of Western Europe, AD 500-1850*.

Smith, P., Hutchison, D., Sterbenz, J. P. G., Sch□ler, M., Fessi, A., Karaliopoulos, M., Lac, C., & Plattner, B. (2011). Network resilience: a systematic approach. *IEEE Communications Magazine*, *49*(7), 88–97.

Snijders, T. A. B. (2001). The statistical evaluation of social network dynamics. *Sociological Methodology*, *31*(1), 361–395.

Snijders, T. A. B., Steglich, C. E. G., & van de Bunt, G. G. (2008). Introduction to actor-based models for network dynamics. *Social Networks*.

Steen, M., & Hansen, G. H. (2018). Barriers to path creation: The case of offshore wind power in Norway. *Economic Geography*, *94*(2), 188–210.

Toth, G., Elekes, Z., Whittle, A., Lee, C., & Kogler, D. F. (2020). *Technology network structure conditions the economic resilience of regions*. Utrecht University, Department of Human Geography and Spatial Planning ….

Trippl, M., Grillitsch, M., & Isaksen, A. (2018). Exogenous sources of regional industrial change: Attraction and absorption of non-local knowledge for new path development. *Progress in Human Geography*, *42*(5), 687–705.

Tsouri, M., Hansen, T., Hanson, J., & Steen, M. (2020). *Knowledge recombination for emerging technological innovations: the case of green shipping*. Utrecht University, Department of Human Geography and Spatial Planning ….

Tsouri, M., & Pegoretti, G. (2020). Structure and resilience of local knowledge networks: the case of the ICT network in Trentino. *Industry and Innovation*, 1–20.

Vermeulen, B., & Pyka, A. (2017). Supraregional relationships and technology development. A spatial agent-based model study. In *Innovation Networks for Regional Development* (pp. 273–290). Springer.

Vermeulen, B., & Pyka, A. (2018). The role of network topology and the spatial distribution and structure of knowledge in regional innovation policy: a calibrated agent-based model study. *Computational Economics*, *52*(3), 773–808.

Vermeulen, B., & Pyka, A. (2014a). Technological progress and effects of (supra) regional innovation and production collaboration. An agent-based model simulation study. *2014 IEEE*

Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr), 357–364.

Vermeulen, B., & Pyka, A. (2014b). The effects of supraregional innovation and production collaboration on technology development in a multiregional world: A spatial agent-based model study. *International Conference on Cellular Automata*, 698–707.

Walker, R., & Storper, M. (1989). *The capitalist imperative: territory, technology and industrial growth* (Vol. 2). Oxford: Basil Blackwell.

Xiao, J., Boschma, R., & Andersson, M. (2018). Industrial diversification in Europe: The differentiated role of relatedness. *Economic Geography*, *94*(5), 514–549.

Yayavaram, S., & Ahuja, G. (2008). Decomposability in knowledge structures and its impact on the usefulness of inventions and knowledge-base malleability. *Administrative Science Quarterly*, *53*(2), 333–362.

Zhuang, E., Chen, G., & Feng, G. (2011). A network model of knowledge accumulation through diffusion and upgrade. *Physica A: Statistical Mechanics and Its Applications*, *390*(13), 2582–2592.