Review Article

Creating Agent-Based Energy Transition Management Models That Can Uncover Profitable Pathways to Climate Change Mitigation

Auke Hoekstra, Maarten Steinbuch, and Geert Verbong

Eindhoven University of Technology, Eindhoven, Netherlands

Correspondence should be addressed to Auke Hoekstra; a.e.hoekstra@tue.nl

Received 6 June 2017; Revised 17 October 2017; Accepted 12 November 2017; Published 28 December 2017

The energy domain is still dominated by equilibrium models that underestimate both the dangers and opportunities related to climate change. In reality, climate and energy systems contain tipping points, feedback loops, and exponential developments. This paper describes how to create realistic energy transition management models: quantitative models that can discover profitable pathways from fossil fuels to renewable energy. We review the literature regarding agent-based economics, disruptive innovation, and transition management and determine the following requirements. Actors must be detailed, heterogeneous, interacting, learning, and strategizing. Technology should be represented as a detailed and heterogeneous portfolio that can develop in a bottom-up manner, using endogenous feedback loops. Assumptions about discount rates and the social cost of carbon should be configurable. The model should contain interactions between the global, national, local, and individual level. A review of modelling techniques shows that equilibrium models are unsuitable and that system dynamics and discrete event simulation are too limited. The agent-based approach is found to be uniquely suited for the complex adaptive sociotechnical systems that must be modelled. But the choice for agent-based models does not mean a rejection of other approaches because they can be accommodated within the agent-based framework. We conclude with practical guidelines.

1. Introduction

Scientists are “95% certain that humans are the main cause of global warming [which] will lead to high to very high risk of severe, widespread and irreversible impacts globally” [1]. They advise “stabilizing temperature increase to below 2°C relative to pre-industrial levels” but note that “this will require an urgent and fundamental departure from business as usual” [1]. Politicians in 194 countries signed the “Paris Agreement” that pledges to keep global warming below the 2°C mark [2]. Anthropogenic greenhouse gases are largely caused by humanity’s use of fossil fuels to power energy intensive machinery. Electricity and heat are responsible for 25%, and it is tightly connected to transportation (which might become electrified) with 14%, industry with 21%, and other energy related greenhouse gasses with 10% [1]. We need models that help us to uncover quick and preferably profitable pathways to a renewable energy system [3].

Since our industrial civilization is built on fossil energy that now permeates every aspect of it, switching to renewables is a fundamental and integral transition that is not easy to model. Renewable energy is also decentralized and intermit-tent with a larger role for prosumers (consumers who are also producers) and demand-response and storage, which makes the entire energy system fundamentally different. Current models cannot cope with the magnitude of this transition and offer ineffective and myopic perspectives. But the reality is less pessimistic: renewable technologies are on track to be cheaper than fossil alternatives [4, 5], even more so if we phase out fossil fuel subsidies [6, 7] and internalize risk calculations [8]. Renewables promise economic growth [9], jobs [10], and energy security [11] while reducing geopolitical tensions [12].

So why not take a more positive approach and try to take advantage of the opportunities that renewables offer?
In this paper we build on theories found in agent-based economics, disruptive innovation, and transition management. We formulate requirements and make practical recommendations for models that can not only make better predictions but also empower us to actively manage the transition to renewable energy. We call them energy transition management models and define them as quantitative models that can discover profitable pathways from fossil fuels to renewable energy.

This paper brings together a wide review of both theory and practice in order to show how such models could be constructed. It consists of five parts:

1. “Clinging to Equilibrium” functions as a problem statement, showing that most current models are unsuitable for transitions.
2. “Theories of Disruption” reviews agent-based economics, disruptive innovation, and transition management.
3. “Modelling Requirements” builds on the theories and focuses on actors, technologies, and carbon pricing.
4. “Modelling Methods” evaluates system dynamics and discrete event simulations before concluding that modelling complex adaptive systems requires agent-based models.
5. “Building a Model” contains practical advice.

2. Clinging to Equilibrium

Currently, most predictions regarding the energy transition come from equilibrium models that are simply not suited to dealing with transitions.

Imagine how hard physics would be if electrons could think (Murray Gell-Mann)

2.1. Equilibrium Models Still Dominate. Economics became a science in the period that most social scientists tried to emulate Newtonian physics, requiring humans to be just as predictable as, say, electrons. Adam Smith’s Wealth of Nations (1776) [13] was inspired by Newton’s Principia. Leon Walras translated Smith’s ideas into “Newtonian” economic models called general equilibrium models [14]. Computing them was first done in 1967 and such models were called Applied General Equilibrium (AGE) models [15]. Since the mid-1980s an implementation called Computable General Equilibrium (CGE) models came to dominate economics [16–18] and their newest incarnation called Dynamic Stochastic General Equilibrium Models (DSGE) dominates economics now [19].

“The concept of equilibrium ... is one of the central pillars of the Great Borrowing from physics ... Voluntary exchange matches up buyers and sellers, prices float until everyone is content, and all markets clear. The similarity to the ideal gas law of physics is not at all accidental” [20]. A generally accepted definition does not exist [21] but an often used one is as follows [18]: “CGE models “describe the allocation of resources in a market economy as the result of the interaction of supply and demand, leading to equilibrium prices. The building blocks of these models are equations representing the behavior of the relevant economic agents: consumers, producers, the government, etc. Each of these agents demands or supplies goods, services and factors of production, as a function of their prices.”

Proponents of CGE argue that the models are complex enough to capture the essential features of an economic situation yet simple enough to be tractable [22]. Other claimed advantages are accounting consistency due to the use of a closed accounting system and an accurate measurement of changes in wealth (as defined in macroeconomics) and they are widely used in economics because they “ensure policy-making is guided by a correct theoretical understanding of how economies function” [23]. They also provide a “solid microeconomic foundation” [18] so “CGE analysis constitutes a powerful scientific method for the comprehensive ex-ante simulation of adjustment effects induced by exogenous policy interference” [24].

According to Fagiolo and Roventini [19]: “at the dawn of 2008 – just before the financial crisis unexpectedly hit – a new consensus emerged: the New Neoclassical Synthesis, ... grounded upon Dynamic Stochastic General Equilibrium (DSGE) models.” Proponents claimed that monetary and even economic policy was finally becoming science [25–27]. Currently most energy models are still largely equilibrium models, for example, the World Energy Model (used by the IEA for its influential “Energy Outlook” series), POLES (used by Enerdata), and PRIMES (used by the European Commission) [28]. Stanton et al. [29] also review the following integrated assessment models using CGE: JAM; Igem; IGSM/EPPA (MIT); SMG; WorldScanc (CPB); ABARE-GT; G-CUBED; MS-MRT; AIM; IMACLIM-R; WIAGEM; MiniCAM; and GIM.

2.2. Equilibrium Models Are Precisely Wrong. It is better to be partly right than precisely wrong [30] and notwithstanding their dominance and mathematical precision, equilibrium models have a long list of problems. They are essentially static: they assume that external shocks can take the system from one equilibrium to another but that the transition itself is irrelevant [17]. Furthermore they are top-down models based on “the holy trinity of rationality, selfishness, and equilibrium” [31]. They assume “rational” actors that can be represented by a couple of “representative” aggregated agents [19]. These “rational” actors only strive for utility maximization. They immediately know the utility of every product and price on the market. They are impervious to status, strategizing, populism, idealism, tribalism, hearsay, or marketing. They do not empathize with future generations (see “discount rate” in this paper) or people in other countries (see “discount rate” in this paper). Furthermore, the actors function in “ideal” markets. This means, among other things, no bankers with perverse incentives, no lobbyists, no political games, no idealists, no monopolistic tendencies, no network effects, no incumbent resistance to change, and so on.

Each of the above simplifications has been falsified, both within economics and by findings from other social sciences like psychology, sociology, and political science [20]. But in equilibrium models they endure, hence the accusation that
2.3. Specific Problems with Energy Models. State-of-the-art integrated assessment models (IAMs) usually embed equilibrium models that treat climate policy as an additional constraint while predicting and factoring in climate-related damage [39]. But there is growing consensus in the literature that this conveys a false sense of control and underestimates both the damage of climate change and profits reaped when implementing renewable energy [8, 40–46]. For more, see the discussion on “Pricing Climate Change” in this paper.

Most models also underestimate the potential of technologies that diverge from the status quo and we want to give one example to illustrate the magnitude of the problem. We will use the photovoltaic (PV) predictions of the World Energy Model (WEM) in the World Energy Outlook (WEO) of the International Energy Agency (IEA). Although it is only a partial equilibrium model, we take it as an example because it is probably the most influential energy model in the policy domain [48].

The output of the WEM is shown in Figure 1. NPS stands for “New Policy Scenario” which assumes that a realistic amount of commitments (e.g., to the Paris accord) are implemented into new policies stimulating PV. The thick black line shows that annual additions of solar have been steeply increasing (meaning new or bigger factories for solar panels). We can see that each WEO accepts this past reality because each new scenario has a higher starting point. However, we also see that the WEO essentially predicts that no new factories will be build. This stagnation of the solar industry is predicted over and over again, in every WEO PV scenario since 2002 [47].

Since the WEM is a proprietary model, it hard to pinpoint the cause of the problem. Johnsen points to erroneously high prices; little learning due to little growth; and assumptions about the absence of flexibility that reduce the value of PV [34]. If that were to continue the price per watt in 2030 would be seven cents instead of one dollar, and although that is probably too optimistic, there is a big difference with IEA expectations. Furthermore, the IEA acknowledges that PV has grown exponentially with—an average—43.3% per year over the last 26 years. However, page 74 shows that WEM NPS predicts linear growth from 228 GW in 2015 to 1800 GW in 2050. A yearly production of well under 50GW would suffice to accommodate this growth. In reality, there was already over 75GW annual production capacity in 2016 and healthy growth is expected in 2017 [50].

This disconnection from reality could be due to, for example, sponsor requirements or mental biases like confirmation bias [51], status quo bias [52], or system justification bias [53] but the way the model works could also be a factor. We would argue that, looking at Tables 1 and 2, most of the energy transition management model requirements that we deduce from the literature are implemented partially or not at all. The result is a model that is unable to envision and leverage the exponential developments in solar energy.

Figure 1: Reality versus IEA predictions: annual PV additions. Data from IEA WEO 2002-2016 NPS and REF scenarios [47].
Table 1: Overview of requirements for energy transition management models.

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Global level with technological innovation through science, R&amp;D, and economies of scale. Also, climate impacts and policies. Ability for rich nations to invest in climate mitigation in poorer nations (e.g., rainforests).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agents are detailed, heterogeneous, and strategizing. They learn in interaction with each other and the environment. Behavior is defined through interviews and surveys. Exceptional individuals and institutions can drive innovation.</td>
<td>National level with energy related policy, subsidy, and taxes. Also, large scale energy production/use, high voltage grids, and mobility patterns. Important level for modelling regime resistance.</td>
</tr>
<tr>
<td>Technology is detailed, disaggregated, decentralized, and validated by domain experts. Agent adoption drives endogenous, possibly exponential, bottom-up feedback loops for, for example, solar, wind, storage, and EVs.</td>
<td>Local level with actors that drive adoption and use of new energy technology. Constrained by spatially modelled physical infrastructure, connecting subsystems like grids, roads, buildings, machines, and people.</td>
</tr>
<tr>
<td>Ability to price externalities differently for different actors. Underlying assumptions are explicit and user adjustable, for example, discount rates, Negishi welfare weights, chance of catastrophe, and value of health/ecosystems.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Characteristics of different types of systems and accompanying simulation techniques.

<table>
<thead>
<tr>
<th>Model characterization</th>
<th>Equilibrium</th>
<th>Dynamic systems</th>
<th>Complex adaptive systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>System creation</td>
<td>Static</td>
<td>Dynamic</td>
<td>Adaptive</td>
</tr>
<tr>
<td>Transition &amp; resistance</td>
<td>Top-down</td>
<td>Top-down</td>
<td>Bottom-up</td>
</tr>
<tr>
<td>Dynamic parts of model</td>
<td>Invisible</td>
<td>Observed as aggregate</td>
<td>Can be traced back to individual agents</td>
</tr>
<tr>
<td>Dominant paradigm</td>
<td>Equilibrium (CGE)</td>
<td>SD and DES</td>
<td>Agent based modelling (ABM)</td>
</tr>
<tr>
<td>Simulated system</td>
<td>No simulation</td>
<td>Predefined system</td>
<td>System emerges during runtime</td>
</tr>
<tr>
<td>Strong feedback loops</td>
<td>No</td>
<td>Yes (especially SD)</td>
<td>Yes</td>
</tr>
<tr>
<td>Spatial awareness</td>
<td>Hard</td>
<td>Yes (especially DES)</td>
<td>Yes</td>
</tr>
<tr>
<td>Heterogeneous actors</td>
<td>Hard</td>
<td>Yes (especially DES)</td>
<td>Yes</td>
</tr>
<tr>
<td>Emergent behavior</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Actors can interact</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Actors can learn</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

3. Theories of Disruption

We have determined that equilibrium theory is not a useful basis for transition management. What technologies are more appropriate? We start with agent-based economics that is quickly growing into a mature field within economics. We expand that with theories on disruptive technological innovation. Finally, we look at transition management: a body of literature with a focus on accelerating the energy transition. Thus, we try to capture the perspectives of economists, engineers, and social/political scientists.

The Mecca of the economist lies in economic biology … But biological conceptions are more complex than those of mechanics (Alfred Marshall, 1907)

3.1. Agent-Based Computational Economics. After the fundamental problems of equilibrium models became clear, some economists started looking for new paradigms. They found inspiration in biology and natural evolution where path dependency is just as important (if not more so) than equilibrium. An overview is found in Nelson [54]. Evolutionary economics was then combined with a new approach from the social sciences that took advantage of modern computing power: agent-based modelling (ABM). The result was a very different, post-Walrasian, branch of economics called Agent-Based Computational Economics (ACE) that uses ABM instead of equilibrium models [55]. ACE is based on the complex adaptive systems paradigm and grows economies from the bottom-up, thereby bridging the divide separating micro- and macroeconomics [56–58]. It extends evolutionary economics in four ways: agents are heterogeneous with
individual memories and large autonomy; new interactions are possible (e.g., predatory or cooperative instead of just based on price and quality); evolutionary pressure acts on individuals instead of entire populations which drives individual behavior; and, finally, the implementation is a virtual economic world in the computer that functions without intervention by the modeller [59]. Balint et al. [39] show that ACE is making an especially pronounced contribution in the areas of energy and climate change.

ABMs are used to find the most effective policies to mitigate the macroeconomic effects of climate change. They show that it is effective in combining an approach using sticks (carbon taxes) with one using carrots (incentives to technology adoption) [60], ABM enables us to reframe climate change from a zero-sum game to a coordination problem offering win-win solutions [61]. Dosi and Nelson show that it is important to endogenize the Schumpeterian diffusion of low-carbon innovation [62]. When ABMs take that into account, they show that subsidizing R&D into carbon-free technologies leads to a swift transition and higher economic growth [63, 64]. Isley et al. [65] show how ABM can be used to compare market-based emission reduction policies with some having surprising and transformative results. Determining storage needed in schemes with more intermittency is also an ideal candidate for ABM studies (e.g., [66]). And when it comes to cap defining policies for financing green innovation ABMs are also proving increasingly useful [67].

ABMs also give us a better picture of the costs of climate change in a networked world [68]. They show how the burden hits parts of the population especially hard [69]. They use catastrophes like the flooding of Louisiana after Hurricane Katrina to show that losses increase nonlinearly due to propagation effects [70]. Propagation effects are also studied globally [71, 72]. Combining costs and mitigation into agent-based integrated assessment models (IAMs) is a new field. One of the first examples of combining complex economies with climate models comes from the LAGOM model family that is able to deal with different timescales for modules and spatially distributed energy production and pollution [73–75] but it is not yet an IAM. The first fully fledged agent-based IAM might be the Dystopian Schumpeter meeting Keynes (DSK) model from Lamperti et al. [76].

Climate negotiation and coalition formation is an interesting new field of study for which ABMs are especially well suited [77–79]. Realistic ABMs that take different perspectives into account produce very different result than equilibrium-based models with rational actors [80] and the way the negotiations are structured have a large impact on the results [81, 82]. Policies that return part of a carbon tax back to conforming companies are found to survive longer [83]. As we will see in the discussion on “Energy Transition Management,” individual (instead of highly aggregated) actors and technologies and new ways of interacting are vital for models that aspire to explore novel energy transition pathways and they can also increase stakeholder participation because the model is a closer representation of the reality the stakeholders recognize.

ABMs are now ubiquitous when modelling the electricity sector [84–86] because of their higher explanatory power in general [87, 88] and better ability to explore the effects of increasing amounts of renewable energy in the mix [89–91]. Multiagent systems can also be used to actually run a decentralized power grid [92, 93].

An application of ABMs in ACE that is closely related to the next section on Disruptive Technological Innovation is predicting technology diffusion [94], especially of renewable technology [95–99]. The agent-heterogeneity that ABM facilitates is a central factor in technology diffusion in general [100], and environmentally benign technology specifically, with “eco warriors” and early adopters functioning as launching customers [101]. While some income inequality speeds up adoption, too much hampers it [102, 103]. The interplay between customers and firms is also important [99]. Finally, the diffusion of knowledge among consumers is crucial for the spread of diffusion (contrary to the rational actor model). An interesting side is that ABM gives a result that we have observed in reality: moral persuasion is ineffective [104] at taking away this barrier while graded eco-labels are very effective [105].

3.2. Disruptive Technological Innovation. Burning organic matter has brought humanity hitherto unfathomable energy [78] and prosperity [79] but now we are making the transition to renewable energy [80]. This will not be an incremental upgrade but a transformative system level change in the Schumpeterian tradition of creative destruction [106] with superior technologies replacing inferior ones [107]. The basic reason is simple: while fossil technology is hardly improving anymore [108], there are increasing [109] or even accelerating [110–112] returns for renewable innovations regarding solar, wind, batteries, and synthetic fuels [113].

This improvement is usually modelled using learning curves. Originally the concept was confined to learning inside factories [114, 115] but they are now deployed in a wider context [116]; many technologies exhibit a relatively predictable exponential trajectory that is a more reliable predictor of its development than mere (expert) opinions because human beings (experts too) have a tendency to assume linear trends.

The most famous example is probably Moore’s law [117], but, in energy, it is solar panels that steal the show. We have seen that their price drops 21% for every doubling of production [118]. This has already resulted in a more than hundredfold price decrease in the past 35 years. The price of wind energy has decreased tenfold in the same time [90]. Lithium batteries for electric vehicles do not go that far back but their price has decreased tenfold in the last 17 years alone [93]. So, the technology advantage is tilting dramatically towards renewable technology. Will this continue?

Availability of energy is not the problem: the sun gives us thousands of times more energy (through radiation) than we need [81, 82] and the first derivative (wind energy) is also abundant [83]. The raw materials we need are not really scarce either and since we can recycle—we are not burning stuff—we could continue with renewable energy until the sun runs out. The price reductions will also continue since we have many, many doublings to go and renewable energy uses much less raw material than fossil energy. For example, wind
needs between 2 grams (airborne wind energy [85]) and 10 grams of material per kWh [86, 87] while a coal fired power plant burns about 300 grams per kWh [88, 89]. And electric vehicles running on this renewable energy use only 400 kg of battery (falling to 100–200 kg) to replace over 15 tons of crude oil over the lifetime of the car [91, 92]. So, if human civilization continues it is likely that fossil technology will be displaced.

However, in many areas renewable technology is still inferior. How do we make sure the disruption does not stall? Christensen et al. have put forward the theory of disruptive technology to explain how an inferior new technology can beat an incumbent one [119–121]. They posit that the new technology might be inferior from the vantage point of mainstream criteria, customers or firms, but for some customers there may be performance oversupply: they get more performance along the traditional criteria than they are willing to pay for. Since consumers often experience decreasing marginal utility [122], performance oversupply could even be the rule. Competing firms can continue to increase performance oversupply as a form of nonprice differentiation [123]. But the performance oversupply makes customers attracted to products that give more attention to price or to new criteria that have so far been considered secondary. Incumbents often react too slow to this shift because they are resource dependent [124] on high-value customers that will not accept the inferior new technology yet.

Adner extends disruptive innovation with a demand based and emergent perspective [125]. He introduces functional thresholds below which a consumer will not consider the product and net utility thresholds that specify the highest price consumers will pay and may be based on the customers internal resources [126], capabilities [127], or human capital [128]. Adner also develops a formal model describing how technology A in market segment X will more easily disrupt market segment Y if there is a large preference overlap between the market segments. Adner states that he would have liked to include asymmetries in firm capabilities, market segment size, and economies of scale.

A final dynamic we must mention is the lock-in to interior technological development trajectories. Arthur points out that this has happened to, for example, our keyboard and electricity grid and the type of nuclear reactors we use [109]. It uses the same dynamic of increasing returns but now has the effect of giving an incumbent an ever-increasing advantage, effectively creating a monopoly for a product (e.g., Google or Facebook) or technology (e.g., gasoline cars). Most important in the context of this paper is humanity’s carbon lock-in towards the use of more fossil fuels [129, 130]. Even though we know it is destroying our habitat, the lock-in of fossil technologies is maintained by subsidies and the absence of carbon pricing [131]. It is also spreading to developing countries [132]. We must be able to model innovative ways to break such lock-ins.

Let us give an example touching upon most of these issues to show what we mean. For over a hundred years the only practical rechargeable battery was a lead acid battery. But when the laptop and mobile phone came along, research into lightweight batteries intensified and soon the lithium battery was born. This enables people to build electric sports cars like the Tesla Roadster that changed the image of electric vehicles for the better.

Tesla went on to produce the model S: an upmarket electric car with good looks, range, and specs. The Model S has now started to displace many conventional car types in the upper car segment. It is inferior in the sense that it has limited range (450 km per charge), takes longer to recharge (about 250 km in 30 min), and offers a less premium finish. But customers—especially people interested in new technology with access to either private parking or a public charger close to their house—liked that it had unparalleled acceleration (faster than the fastest production sports car) and was silent and clean [133]. Once they owned the vehicle they noticed how they loved charging at home (which they did more than 90% of the time) and how the range was not really an issue [133].

The price of the batteries was still too high to make a mid-level car but the success of the Model S enabled Tesla to build a Gigafactory for batteries (together with Panasonic) that brought prices down [134]. Government subsidies and charging infrastructure are currently also vital [135]. But prognoses show that electric vehicles will soon be cheaper to own due to falling battery prices, low energy use, and low maintenance [135, 136] after which displacement will continue without subsidies. This example contains a number of dynamics that an energy transition management model should be able to capture.

3.3. Transition Management. In our view, transition management enriches theories on agent-based economics and disruptive innovation with political and sociological insights. It also specifically addresses lock-in of fossil technologies. Energy transition management was born from the observation that transitions to new technologies often fail to materialize, even if they are beneficial to society. In order to change this, transition management researchers aim to “develop appraisal and valuation techniques that could inform a choice between different technologies” from the vantage point of societal benefit [137]. Later this technology focus was extended to “interconnected systems of artefacts, institutions, rules and norms” and especially on how to make systems transition in response to environmental signals [138].

Recent overviews of the field are given by Sengers et al. [139], Li et al. [140], and Holz et al. [141]. According to Li et al.: “Today’s most influential body of innovation-focused transition research originates in the Netherlands, and is often called the “Dutch approach”. Approaches that descended from the Dutch school are transition management (TM), strategic niche management (SNM), technological innovation systems (TIS), and the multi-level perspective (MLP).” These approaches are particularly suited for investigating sociotechnical transitions in the energy supply, buildings, and transport sectors, as they focus on means of supplanting the incumbent system with radical alternatives, disruption of the status quo, and the initiation of rapid change.

If this paper refers to transition management, this is meant to include SNM, MLP, and TIS. All these approaches focus on regime resistance from incumbents and how
protected niches can function as incubators for beneficial innovation [142–144]. SNM [145, 146] draws on evolutionary economics [147, 148] and had high ambitions (although the originators concede that “we were certainly over-optimistic about the potential of SNM as a tool for transition” [149]). MLP [150] focuses on the fact that change is required at multiple levels at the same time. As the name implies, the study of technological innovation systems (TIS) [151] is focused on technology instead of on the energy transition as a whole [152, 153]. However, since the energy transition is driven in part by these technological innovation systems, their integration into transition management yields powerful synergies [154, 155].

Transition management sets out to explore radically new ways to meet societal needs for, for example, energy or mobility [156, 157] and draws heavily on the ideas of Malina and Kauffman [158] regarding complexity [159]. It combines many of the aspects of the other approaches and gives special focus to individuals and organizations that are aligned to both the innovative technology and the societal goal (e.g., sustainability) since it is them (and not members of the incumbent regime) that will drive the change [160, 161]. However, this requires “multiple explicit actors with differentiated selection criteria and behavioral parameters that possess agency to shape transitions” [140]. An intriguing new idea is to use the transition management perspective in conjunction with crises to create “game changers” [162].

Currently, transition management models are usually qualitative in nature. Drawing on Holtz et al. [141] we conclude that quantified transition management models are needed because they improve internal consistency [163–165] and enable validation of theory [166–168]. Models can also enhance participatory processes [169, 170] while challenging incorrect narratives [171] and bringing relevant factors into scope [172, 173]. Finally, quantified models would increase the use of transition management by policy-makers which is needed because currently it is used only rarely [141].

4. Modelling Requirements

In this section we present the requirements for energy transition management modelling that we deduce from the theory in the previous section first about modelling actors, then about modelling technologies, and finally about carbon pricing.

4.1. Requirements for Actors

4.1.1. Detailed and Heterogeneous Actors. According to transition management (esp. MLP), there is a range of transition pathways. Geels et al. [174–176] identify the following:

(1) Transformation: incumbents adjust some regime rules under pressure of, for example, social movements, for example, the adoption of carbon capture and storage (CCS).

(2) Technological substitution: novelties are developed before incumbents act and they are replaced within the regime, for example, electric vehicles and large scale solar.

(3) Reconfiguration: a range of novelties is adopted by incumbents and this deeply changes the regime, for example, self-driving vehicles, supergrids, and smart grids.

(4) De- and realignment: landscape pressure, uncertainty, and multiple novelties lead to a new regime, for example, solar plus storage replaces centralized energy systems with autarkic micro grids.

We find that models that strongly aggregate and simplify actors mostly capture the first type of pathway. Transition management can capture all pathways by viewing transitions as driven by specific innovators that are part of a niche and resisted by specific incumbents that are part of a regime. But this requires heterogeneous actors that are modelled in detail. Furthermore, the actors possess agency. The transition management view of agency could be characterized as follows: “Never doubt that a small group of thoughtful committed citizens can change the world. Indeed, it is the only thing that ever has” [177]. In transition management it is the actor that goes against the grain that is the nucleus for change (together with technological breakthroughs), whether at the global level (e.g., Henry Ford, Steve Jobs, or Elon Musk) or at the local level (e.g., a community organizer or buyer of solar panels and electric vehicles). At first the innovator is part of a niche. A niche is created when the innovator is joined by a group of actors that are open to the innovation. After a while the bottom-up innovation can grow out of the niche to become a new incumbent. The implication of this idea is that change is caused by (groups of) actors that diverge from the mainstream. So actors especially that might act as innovators or early adopters must be modelled in detail.

But not all actors are individuals. Important actors in the energy transition are governments, knowledge institutions, corporations, and NGOs [178, 179].

Different stages of the transition see different innovators as its drivers. For example, solar cells were first developed for satellites, then adopted for buoys and ships, then adopted by households in sunny countries with high electricity prices, and only now are they becoming the cheapest form of energy in wholesale installations. In each phase the innovator was a different actor.

The importance of bottom-up modelling can also be illustrated by the way we have adopted new technologies like PCs, the Internet, mobile phones, and digital cameras. Every one of these trends surprised incumbents [180] and such trends are typically underpredicted by economic (top-down) models [181]. A similar underestimation is currently happening regarding PV, EV, and wind. If we switched our perspective to the actors actually adopting the new technology—because they perceive a benefit—this would remove the problem of underestimating change.

A way to gather the relevant actors and behaviors suitable to transition management could be to conduct interviews, focus groups, and surveys among human actors and organizations that could be instigators for change and to make sure that the diversity of their behaviors and interactions is captured in the model. An example would be smart-grid users [182]. Representation and aggregation would only be
acceptable if a valid line of reasoning suggests it does not diminish the ability of nonstandard actors to instigate or resist change. When the model is ready the modeller should preferably return to the flesh and blood versions of the actors represented in the model and gather their feedback on how well they are represented.

### 4.1.2. Interacting and Learning Actors

A characteristic of particular interest is how actors interact with each other and the physical infrastructure.

In sociotechnical systems interrelations are important and determine, for example, the success of innovative entrepreneurs [183]. Innovation is spread by actors interacting and learning from each other [184, 185]. Interactions with the physical environment are just as important. For example, people cannot put solar panels on their roof if they do not have a roof and for some countries the adoption of land based wind power is easy (like the USA) while for others off-shore wind is more logical (like the UK and the Netherlands). In transportation the importance of geographical locations and infrastructure is even more obvious. For example, adoption of electric vehicles is less for actors that depend on public charging infrastructure that does not materialize. Finally, actors can learn. An example of this is that people seem to like electric vehicles better with experience [186].

An approach we would recommend is to copy a representative set of environments of interest (e.g., neighborhoods) in which actors can “live” and interact in a realistic fashion. For example, when modelling electric vehicle charging you would model houses, parking places, and charging stations and give actors realistic buying, charging, and driving behaviors. Realistic computerized neighborhoods can provide a scaffold for scenarios in which the modeller is forced to “face reality” and it is beneficial in interactions with domain experts and policy-makers because it provides them a recognizable interface.

### 4.1.3. Actors with Foresight and Strategizing

Transition management actors need foresight and strategy. For example, Karl Benz had the foresight to develop the internal combustion engine (but would have gone bankrupt if his wife had not had the foresight that he needed buyers so she kidnapped the prototype to make the world’s first "long distance car trip" [187]). Henry Ford faced customers who had no idea what an automobile was but he foresaw the age of the automobile. Elon Musk has claimed at several occasions that he started Tesla because he thought he could prove that electric cars were cool, even though he considered the chances of success for the company to be less than 10%.

A more statistically relevant example is investment in solar panels. The investment is now more than hundreds of billions of euros which is only rational if the investors have considered a longer time frame and exponential developments.

Regime actors can also use foresight to stave off challenges to the status quo, for example, by sponsoring scientists that support their view and by paying lobbyists and politicians that help to keep/make the playing field uneven.

### 4.2. Requirements for Technologies

#### 4.2.1. Realistic, Disaggregated with Endogenous Learning

Transition management models require a “disaggregated portfolio of technology options with different price and performance characteristics that function within operational boundaries and face resource constraints” [140]. In practical terms, each individual technology should be modelled as a separate bottom-up contender for adoption, including all the characteristics determining its adoption such as the space practically available and the measures needed to overcome intermittency (like storage and smart charging of electric vehicles).

Especially important is the inclusion of learning curves as detailed in the previous theory section. Learning curves can be exogenous or endogenous. An example is the analysis of learning curves of batteries by Nykvist and Nilsson [188]. They present an exogenous feedback loop where the costs of battery production are declining by 8–14% per year. Then his endogenous feedback loop shows battery prices declining at 6–9% per doubling of production capacity. The second curve is called an endogenous feedback loop because it feeds on itself within the model: an increase in battery sales would lead to a corresponding decrease in the price of batteries which could lead to an increase in battery sales, and so on. Models that assume endogenous technical change frequently recommend much more aggressive carbon abatement policies and see a much quicker and less costly replacement of fossil fuels by sustainable alternatives. A review of 25 well known models concludes that the impact is significant [189].

Feedback loops can be negative (dampening the reaction of the system) or positive (quickly taking the system to new states). Positive feedback loops are important because “they generate a much richer variety of trajectories which the system may follow” [190]. One objective of transition management models is to enable policy-makers to spot and then accelerate positive feedback they deem beneficial, for example, by protecting niches (when they have the potential to provide powerful positive feedback) from regime resistance until they are self-propelling.

Another aspect is the facilitation of bottom-up adoption that leads to the emergence of new pathways. Top-down models “tend to be more pessimistic than bottom-up models about the costs of energy policies. This is related to the difference in the scope and potential for energy efficiency improvements as well as the treatment of technological change” [181]. Verbong and Geels [175] describe how different pathways can lead to different solutions, for example, either to micro grids, smart grids, or supergrids [176, 191]. By allowing new pathways to emerge in a bottom-up manner, transition models allow us to find better ways forward.

#### 4.2.2. Use of Real World Technological Experts

When translating the real world into a model it is valuable to do it in such a way that the model is easy to understand for real world experts. Then modellers (e.g., economists) and experts (e.g., engineers) could work together. Let us take PV as an example.

PV production is comprised of many subprocesses with different rates of learning. In the case of PV, traditional
cells benefit from our ability to slice ever thinner silicon wafers in an ever more efficient fashion while thin film technologies benefit from breakthroughs in printing cells on a flexible substrate. Experts could help us judge those different learning curves. They could also double-check them based on predicted raw material usage. In the case of PV, a silicon wafer that makes up a solar cell has a minimum thickness and can never become cheaper than the silicon it is made off. The fragile wafers also require a rigid (e.g., glass) encapsulation of a certain thickness. That is why many experts predict that thin flexible film will win out over silicon wafers in the long run. But things can change: when the price of silicon skyrocketed a few years ago the PV industry developed newer and simpler processes specifically for PV (instead of borrowing from the processes used to make CPUs) that now produce solar silicon significantly cheaper than before. So, we have to keep monitoring and updating our models. Using models that closely resemble reality makes it possible to interact with domain experts that can help us with that.

This leads us to R&D. One could argue that Einstein's Nobel prize winning explanation of the photoelectric effect paved the way and that our current computer models of quantum gaps and the way we apply machine learning to the simulation of digital composites are an important factor in the constant PV breakthroughs. R&D has the advantage of being relatively cheap (compared to scaling up production) but the disadvantage of being relatively slow (especially fundamental research at universities). Thus, investing in more PV R&D might be a very cost-effective way to increase learning rates in the long run but it requires foresight. Experts could help us to incorporate the possible impact of R&D in our models.

But although domain experts are vital, there is a catch. Experts are just as critical of paradigm changing feedback loops as ordinary humans and can have a stronger than average status quo bias [52] and system justification bias [53]. Maybe that is why science is said to progress "one funeral at the time" and why the science fiction writer Clarke proclaimed his "first law" to be as follows: "When a distinguished scientist states that something is possible, he is almost certainly right. When he states that something is impossible, he is very probably wrong" [192]. These biases can combine with groupthink that exacerbates confirmation bias [51] and are in turn exacerbated by shared information bias [193], especially when the group is seeking consensus [193]. In short: experts tend to be conservative, especially in groups.

For this reason, the model should follow the knowledge of experts when capturing the dynamics and constraints of the technology but predictions about the future should be based on where the model leads us, and that will probably surprise both us and the experts.

4.2.3. Externalities of Imperfect Markets. Externalities are costs or benefits (usually costs) that are not internalized in the transaction under review. For example, we know that burning coal fuel and eating meat contributes to climate change but the costs of climate change are not incorporated in the price of coal or meat. This means that producers can offer fossil fuels and meat at low prices and that society as a whole picks up the tab.

Governments could remove these market imperfections by levying taxes, using a principle often called the polluter pays. But this causes regime resistance, for example, from the producers and consumers of fossil fuels and meat. We also face a prisoner's dilemma in the sense that industries in the first country to adopt taxes would find it hard to compete with industries from other countries that did not adopt the taxes. So, taxes that are beneficial on a global level can be detrimental on the national level if some countries opt out.

Transition management models looking for beneficial pathways should assume neither a continuing status quo nor perfect markets. Instead they should focus on the overall costs and benefits of realistic interventions. These costs should be attributed to individual actors on different levels (e.g., the planet, individual nations, and individual producers and consumers) so that resistance and gain are visible on the level of the relevant decision-makers. This holds true for all parts of a transition management model but is especially relevant when modelling technological pathways.

For example, a pathway that assumes erasing subsidies for meat or forcing people not to eat it is highly unrealistic. But subsidizing tasty meat replacement products might be achievable, especially in the early stages when production volume is still low. Investing in R&D towards meat replacement might be even more cost-effective and have a higher chance of collecting the ROI locally. Such a pathway is especially promising if the meat replacement technology has the potential to become cheaper and tastier through the existence of a positive feedback loop. Then it would be able to replace regular meat, even in an imperfect market where the costs to society are not internalized.

Thus, well-crafted transition management models have the potential to shift the dialogue from ideological trench warfare around bland win-lose solutions to an intellectual game where the goal is to discover an achievable win-win pathway. And because the pathway is constructed in a bottom-up manner using recognizable building blocks (instead of impenetrable mathematics) it could also be used to create broader support for the investment opportunities that are involved. It might even convince the meat producer to enter the meat replacement business.

4.3. Requirements for Carbon Pricing. The last set of requirements for models that guide the transition to renewable energy pertain to pricing the damage done by fossil technologies. This section should make the transition management modeller aware of the most important assumptions regarding what is often called the social cost of carbon (SCC).

We propose transition management models should be able to use a different SCC, depending on the actor, for example, a high price for actors focused on the state of the art in science, a medium price for economists, a low price for policy-makers based on what is politically palatable, and a negligible price for companies that have the power to lobby for exemptions. Incidentally this could also be a realistic way to model regime resistance. In an ideal model (not an ideal world), discount rates should be context sensitive and all assumptions should be user configurable. We will illustrate the importance of these points below.
4.3.1. Economists versus Climate Scientists. A meta-analysis of 800 estimates showed a “bias-corrected” SCC of between 0 and 134 dollars while “uncorrected” prices are often much higher [194]. But narrowing it down shows that most economists agree that 37 dollars is a safe lower bound and this is what many policy-makers strive for [195]. The actual price on the EU ETS market at the time of writing was 4 dollars. Why this is important can be illustrated by a simple example. Producing one kWh of electricity from a coal fired power plant emits around one kg of CO$_2$ [196]. A CO$_2$ price of 37 dollars per ton would thus lead to a SCC of 3.7 cents per kWh. This would almost double the wholesale price of electricity from coal.

And, actually, the economists are probably underestimating the SCC. Nordhaus (possibly the most influential economist in the field) noticed a deep divide between climate scientists and economists regarding the SCC. He conducted a survey [197] and concluded that “At one extreme are mainstream economists who view the prospect of greenhouse warming with little concern, confident that human societies will adapt handily to such changes. At the other extreme, natural scientists worry about irreversible impacts on natural systems. They warn of unpredictable extreme events, such as shifting ocean currents or migrating monsoons. [Concerning damage] natural scientists’ estimates were 20 to 30 times higher.”

Climate scientists include tipping points and other non-linear relationships in their analysis [198] which creates a discontinued risk profile akin to what is used to calculate the fire insurance for a house [8, 40]. This, combined with the limited availability of data, leads statisticians like Ikefuji et al. to apply Weitzman’s “Dismal Theorem” [199] which gives the risk profile a relatively “fat tail” like a Student’s $t$ distribution [200]. The result is an SCC of hundreds of dollars [201–203]. Economists like Nordhaus dispute the applicability of the Dismal Theorem [204] and treat climate change as a continuous cost subject to cost-benefit analysis (e.g., the price of electricity) to arrive at a much lower SCC. So, the question becomes: who do we trust to estimate the risks of climate change: climate scientists or economists? If we trust the climate scientists or even if we choose a middle ground between them and the economists, the SCC would be so high that fossil fuels would instantly be prohibitively expensive.

4.3.2. Discounting the Future, Hardwiring Inequality, and Pricing the Priceless. The extent to which future damage is taken into account depends on the discount rate [205]. The impact is most easily demonstrated by an example. Assume a damage of a hundred euros to people living one hundred years from now. How much should we invest today to prevent that future damage? Many important economists [206–208] argue that we should not burden future generations with our excesses. They thus argue for a discount rate of 0% and want to invest a hundred euros now. Stern [8] would advocate for a compromise at 1.4% (higher than what governments pay for a loan but much lower than what banks ask of a company) leading to an investment of twenty-five euros. Nordhaus [209] argues that this is an ordinary market transaction and that a prudent rate is 4.3%, which would mean investing little more than one dollar. Some consensus seems to be building towards distinguishing between the interest rate on risk capital and a social-welfare-equivalent (e.g., the 1% rate on obligations) for use in SCC [210, 211]. This last approach would validate Stern’s SCC of roughly two hundred euros. If you look at the previous illustration you will see that this would make electricity from coal five times more expensive. We concur with Kaplow et al. [212] that the discount rate is too important to leave to the modeller because this one assumption can completely alter the output of the model. We contend that the discount rate should be user configurable and in the case of static output (like a paper report) there should at least be a sensitivity analysis present that shows readers what happens when different discount rates are applied.

Another obscure part of SCC calculations is Negishi welfare weights. This mechanism basically assures that money cannot leave rich countries to be put to work in poor countries where it would save more lives [213]. Although this corresponds to political reality it is good that the user of the model is aware of this, for example, because it can alert her or him to cost-effective win-win solutions that are now routinely forgotten like protecting ecosystems in their country of origin.

The problem of pricing goes beyond the discounting of poor and future people and ignoring catastrophes. Economics primarily deals with putting a price on goods humans produce and consume and in those cases pricing is relatively clear. But how to put a price on the destruction of ecosystems, loss of biodiversity, depletion of aquifers, desertification of fertile ground, and the relocation of people? Are pandas worth more than bees? Are European woods worth more than rainforests? Must the calculation be based on tourist industry earnings or can things have an intrinsic dignity as Immanuel Kant proposed? How many km$^2$ of rainforest have to disappear before we admit to a link between biofuels and land use change? These are fundamental questions that relate to norms and values that can only be judged subjectively but are currently decided by modellers that hide them from the policy discussion.

Table 1 summarizes the requirements that have been established so far in the paper.

5. Modelling Methods

So far, we have denounced equilibrium models, reviewed theories, and deduced requirements. Now it is time to describe how better models actually work.

There are myriad ways to characterize what is happening in energy modelling [28, 214–216]. Even the classification of the underlying simulation methods is diverse [217, 218]. We will take the perspective of Borschchev and others [218–220] to create a narrative describing how modelling techniques for sociotechnical systems has become more realistic over time with the help of computers.

5.1. From Static to Dynamic Systems. Energy transition management models focus on "transition pathway dynamics: assessment of normative goals; radical alternatives to incumbent status quo technology or behavior options; time
horizons sufficient for exploring long-term sociotechnical change; and path dependencies” [140]. We saw that equilibrium models cannot produce such pathways. System dynamics (SD) offers a big step forward. SD models consist of an ordered set of differential and algebraic equations that model stocks (also called states) connected by flows (signals: inputs and outputs) [221]. This approach became practical after the advent of computers. Originally SD was implemented on analog computers where op-amps represented the differential equations but it soon turned to digital computers [222].

As described under technologies, some of the most interesting dynamics in transition management occur when the model contains endogenous learning curves for technology. For example, the adoption of PV leads to economies of scale that make it cheaper, which leads to more adoption, and so on. In static modelling this is a problem because it can lead to so-called algebraic loops for which an analytical solution can only be found using nonlinear optimization. This problem often occurs in energy models [223]. If time derivatives are added to the model, as is the case with SD models, the feedback mechanisms can easily be incorporated.

SD models are widely used in biology, chemistry, and engineering [224] with notable examples in mechatronics [225], multibody vehicle dynamics [226], and the spread of disease [227] and also in the social sciences with notable examples being marketing [228]; policy analysis [229]; project management [230]; and learning [231, 232]. Many people know SD from the study “Limits of Growth” that drew attention to the ecological problems caused by humanities ever-increasing use of natural resources [233, 234]. SD has also been widely applied to energy models [235–241]. Within transition management it is mainly used in the context of technical innovation systems [152, 155, 242–245]. SD has longer been used in biology since CAS was first adopted. One of the founders of SD—sketched a picture in which SD had to leave its current plateau and climb new mountains. He was especially disappointed by the lack of use in economics and the public sector [246].

A relevant characteristic for transition management is that SD models contain stocks. For example, in a country the existence of roads, houses, fertile ground, and aquifers that provide sweet water can be modelled as stocks that represent value. Even soft factors like a shared culture and trust in government and democracy are stocks vital to wealth that could in theory be modelled. Monitoring stocks is especially important for environmental models. For example, the build-up of pollutants or greenhouse gasses, the depletion of aquifers, or the effects of land use change due to the use of biomass.

After adding the dynamic of time, we should add the dynamic of space. In energy systems the spatial distribution of energy production, transportation, and use is highly relevant [139, 247–250]. For example, electricity is produced by power plants, windmills, and solar panels at certain locations, transported by a geographically distributed power grid that poses practical costs and constraints and used by agent and machines like residents and electric vehicles that can all be a bit different.

If this was all the dynamism we were after we could use discrete event simulations (DES). As the name implies, DES processes the simulation one event at a time [251] and it is well known for things like complex and spatially distributed queuing problems and Goldratt’s theory of constraints [252]. Simply put: if you want to model an assembly line or airport, DES is probably your first choice.

DES is more recent than SD because it is even more dependent on computers: calculating every event and constantly updating the state and position of every product is computationally intensive, especially since models can no longer be reduced.

In most developed countries the geographic layout of these subsystems has been described and is now publicly available in the form of shapefiles. We can create a spatial environment with many layers that connect events in space. Let us give an example. Electric vehicles can use the road layer to determine how long they have to travel. The building layer tells us where the residents live and we choose a parking place close to that location. The grid layer gets activated as soon as the electric vehicle starts charging, and so on. In this way a many-layered GIS file enables us to create complex, realistic, integral models with relative ease.

Realistic geographic environments are also valuable when the modeller interacts with domain experts or policy-makers. Running the model on a recognizable geographical substrate is visually stimulating, makes it easier to understand what is happening in the model, supports the detection of unrealistic behaviors, and creates trust in the model results.

### 5.2. From Dynamic to Complex Adaptive Systems

But this is still not enough. Researchers now know that many domains, and certainly energy and mobility, need to be studied as complex adaptive systems (CAS) [253]. The first thing to understand about CAS is that the system itself is constantly adapting. You must model the system in a bottom-up manner using autonomous actors and their behaviors in space and time, and as the actors interact with each other and their physical environment, systemic order emerges. (This cannot be modelled in system dynamics or discrete event simulation [254].) Although top-down control of this order is absent, it is adaptive in the sense that individual and collective behavior can mutate and self-organize in order to create a system that can be more robust [255] and resilient [256] than comparable systems with top-down control. The Internet is a good example [257].

Biological was among the early adopters since CAS was able to portray ecologies with complex behavior emerging from simple rules. Wilson used it to popularize sociobiology [258], Dawkins to explain selfish genes [259], and Goodwin and Saunders to improve developmental biology [260]. In chemistry Prigogine used CAS to increase our understanding of dissipative structures and irreversibility [261]. In particle physics, famous adherents were Feynman and Gell-Mann [262]. CAS was even used successfully in philosophy [263]. Holland and Reitman used CAS to create genetic algorithms [264] while Holland also wrote one of the best introductions to CAS [265, 266] and created the often used definition of “systems that have a large number of components, often called agents, that interact and adapt or learn” [267].
Although early ideas pointing to CAS can be found in sociology [268] the uptake in the social sciences was slow [269] and it took until 1996 for it to be applied to social systems by Epstein and Axtell [270] and Axelrod [271, 272]. In economics Tesfatsion played a pivotal role from 1998 onwards [59] while 2006 was a “golden year” with the handbook on Agent-Based Computational Economics [55] and powerful arguments by Foster [58] and Beinocker [273].

Properties of complex systems that emerge from the interplay between underlying processes [274] are highly relevant to transition models [174]. Transition management, SNM, and MLP view the transition to renewable energy and sustainable transportation as “a radical, structural change of a societal (sub)system that is the result of a coevolution of economic, cultural, technological, ecological, and institutional developments at different scale levels” [159] which require a complex adaptive multilevel multidimensional systems approach [158, 275, 276].

Complex adaptive systems can change their ontology. In transition management this is indicated by the term deep uncertainty [277–279]. Examples are the introduction of a radically new technology (e.g., electric vehicles) that replaces an old technology or the rise of a new kind of actor (e.g., prosumers). This is something that is relatively easy to model in agent-based simulation but (almost) impossible to do using other methods.

5.3. Agent-Based Modelling. We already spoke at length about Agent-Based Computational Economics. But how do agent-based models work exactly?

First it is good to appreciate that complex adaptive systems are counterintuitive for primate brains due to non-linearity, feedback, time delays, and interdependencies [231]. Various empirical studies have demonstrated the severity of the problem [280–284]. Fortunately, there are ways to simulate complex adaptive systems using a computer [270, 285, 286] so now we can begin to understand them [287].

Models that accomplish this are called agent-based models (ABMs). Although one could theoretically trace ABM back to the Von Neumann machine, the first real ABM (using coins on a board) was Schelling’s segregation model from 1969 [288]. After some experiments in the 80s by, among others, Axelrod and Reynolds it was Holland and Miller who used the term “agent” for the first time in 1991 [289] and it was only in 1996 that Epstein and Axtell [270] and Axelrod [271] explicitly introduced ABMs to the social sciences. Axelrod and Epstein both called it a third way of doing science [272, 290]: generative and bottom-up and distinct from the usual inductive or deductive approaches. In recent years an increasing number of reviews note their promise for advancing the social sciences [291, 292]. And, as we already saw, ABM is increasingly used in economics.

ABM is now well accepted in empirical social research [293] and the preferred modelling method for sociotechnical systems [294, 295]. Recent reviews note that ABM is now widely used, for example, in health [296] and more specifically oncology [297]; epidemiology [298]; chronic [299] and noncommunicable diseases [300]. Other reviews look at their use in organizational science [301]; emergency response [302]; land use [303, 304]; manufacturing [305]; ecosystem management [306]; and marketing [228, 307].

Energy systems are a particularly fertile area for ABM with reviews in the fields of energy modelling in general [253, 308–310], smart grids [311–313], electricity markets [314–316], distributed generation [317], and transportation [318–321]. A recent review of transition management models found one DES, six SDs, and seven ABMs [140] which is remarkable considering that in the modelling world at large DES and SD are still much more common than ABM.

The reasoning of these agents is becoming increasingly complex. In most cases we find task oriented actors to be the most efficient and easy to understand implementation but many authors advocate explicitly programming different agent states [219] and some like to rely on the belief-intent-desire model [322–324]. Niches, incumbents, and regime resistance can be modelled with relative ease but should ideally not be represented by aggregates but by the individual actors that they consist of in reality. It is also relatively easy to model disaggregate technologies (including endogenous learning curves) and radical innovation, driven by entrepreneurs and policy-makers with strategy and foresight but slowed down with regime actors possessing the same.

5.4. Multilevel Agent-Based Models. Finally transition management models need to be multilevel. In the multilevel perspective (MLP) that we already described, the upper level is often called the landscape, the intermediate level the regime, and the bottom level the niche [150, 191, 325, 326].

On the global level we have issues like climate change, the Paris treaty, and technological developments. The latter are of special interest to transition management models, for example, price developments in silicon solar cells, off-shore wind, heat pumps, NCM batteries, and electric vehicles. We also have R&D breakthroughs about to get out of the laboratory like metal-air batteries, solar fuels, airborne wind energy, and the application of “wonder materials” like graphene. All these developments have the power to shape what happens at other levels by making renewable technologies more attractive to actors on lower levels.

There are many possible intermediate levels, for example, a cooperation that together buys solar panels and heat pumps; or an industry site or municipality that decides to develop an integral energy plan for dealing with heat, electricity, a wind park, and electric mobility, making the combined business case much easier; or a country that decides to promote EVs and change rules and regulations to enable self-driving vehicles.

On the bottom level we have, for example, small companies, households, and persons that can develop, buy, and use solar panels, heat pumps, and electric vehicles. They can also invest or divest in technologies. Their decisions drive the energy transition on all levels but are influenced by developments on higher levels.

We could theoretically model every actor in the world but that would be impossible to program and compute. By choosing a multilevel approach we can choose representative actors on every level and scale, for example, with one actor per global policy and technology, a couple of countries
and municipalities on the intermediate level, and maybe a thousand individuals on the bottom level.

In Table 2 we give an overview of the types of systems and their characteristics as discussed in this paper.

6. Building a Model

We have shown the requirements and methods needed for transition models, but how do you actually build them? This question might appear too practical for a scientific paper, until you realize how little valuable advice is available in the literature and how important this step is for the successful development of a model. This chapter aims to make the knowledge gap in the literature a little smaller.

6.1. Learn from Professional Programmers. Every scientist has a habit of undertaking rigorous literature study but the tools and development trajectories are often chosen less rigorously. We think that, in this respect, modellers should learn from people who are able to work in large teams to develop models containing millions of lines of code: professional programmers. There are strong similarities between object oriented (OO) programming and ABM and it is no coincidence that the dozens of ABM frameworks we know are all written in OO languages. Terms in an OO language like Java often have an equivalent in ABM, for example, object/agent, class/species, and method/action [327]. That is fortunate because OO is now the de facto standard in large software projects because it makes code manageable and does not create “spaghetti code” [328]. We would suggest that, similarly, ABM is optimal for most modelling challenges because it does not produce “spaghetti models.” So, ABM might be the right approach not only for modelling complex adaptive systems but also for other large modelling projects. Using ABM as an organizing principle akin to OO makes it easier to divide the model into independent agents (akin to objects) and to hide their inner workings (using encapsulation). One could even encapsulate entire models. Agents can be hotlinked and used in cosimulations.

Choosing your agents is similar to choosing your objects in object oriented programming. The right choice depends on the modelling situation and/or question. However, we feel we can provide some general guidelines, especially since we advocate the use of ABM as a container for other modelling approaches. From a modelling perspective agents should make independent decisions [271, 329, 330]. Humans (e.g., consumers and institutions or organizations (e.g., households, governments, or energy producers) are logical choices. But agents do not have to be self-conscious. For example, an electric power line could be an agent that “decides” when it is overloaded, thus causing a blackout. From a programming perspective, agents should be viewed as chunks of code that can be developed independently. Using OO principles such agents should have well defined interfaces that are used to initiate behaviors. In this way new agents can be added to the model without impacting the other code in the model.

Of course, it is important to choose a good programming framework and because agent-based modelling (ABM) is such a recent phenomenon, those frameworks are still immature. There are literally dozens of competing ABM frameworks in which the modeller is basically programming directly in an object oriented (OO) language like Java [331]. However, we found that using a dedicated language developed specifically for ABM can greatly enhance productivity [332] (we estimate a factor of ten) and make ABM more accessible to modellers without a computer science background.

A language developed for a specific application is called a “domain specific language.” MATLAB and R are examples of tools (but not ABM tools) that use a domain specific language. The best-known framework using a language especially developed for ABM is called Netlogo. However, we chose the GAMA-platform because it combines the ease of use of Netlogo with the maturity of Java (e.g., building on Eclipse). It also has advanced spatial capabilities (e.g., read and “agentify” shapefiles). Like Netlogo it is free and open source which enables users to extend the domain specific language. Some ABM frameworks are specialized even further (e.g., for smart grids [333]) but we would warn against their use because it inhibits the development of integral models [176].

A key lesson from open source projects is that making the underlying code directly downloadable provides much better error detection and makes it much easier to collaborate on models. For scientific projects and public policy, especially in the important and often politicized case of energy models, the transparency provided by open source models might be even more important [334]. We would go as far as to say that downloadable models are like peer-review for the 21st century. Combining agent-based and their inherent modularity with open source also makes it much easier to create large integrative models. It is worrisome that policy recommendations are often provided by consulting firms that do not disclose their funding or provide their underlying model. If transition management modellers are serious about unleashing bottom-up change and governing the energy transition [335] they should lead by example and share their code online.

6.2. Create Hybrid Models When Needed. Many authors point to the added capabilities from ABM over equation based models [298, 336] and more specifically equilibrium models [20, 55, 217, 337]. Simply put: complex systems can be implemented in ABMs but not in equilibrium models. However, it is possible to include equilibrium models in ABMs. That it is possible does not mean it is advisable and making an ABM that expresses the equilibrium paradigm is probably just a bad idea [19]. However, comodelling simulations are increasingly common [338, 339] and we could imagine scenarios where equilibrium driven agents are part of a larger complex adaptive model [55, 337, 340].

Implementing the stocks and flows of system dynamics (SD) in an agent-based model (ABM) is much more straightforward: simply create an agent for every stock and flow. In most cases you could also model just the stocks and include the flows as behaviors of the agent minding the stock. Heterogeneity can now be added by splitting up stocks into separate agents. For example, instead of one agent representing a shoal of fishes we might create separate agents that represent a single fish. Some fishes might then
be different than others and might be differently processed (e.g., by releasing accidentally netted dolphins). Even more realism can be added by giving each agent a unique location in space. For example, instead of one “flow agent” signifying “fishing” the model could contain multiple agents signifying “fishing boats.” The amount of fish caught could depend on the concentration of fish in that part of the ocean. One application where SD and ABM can be combined is integrated assessments [341, 342].

Similarly implementing discrete event stimulation (DES) in an ABM is easy. Both are inherently location based and can work with maps on which the entities in the simulation have a clearly defined position. In DES the events are implemented using active entities that have a fixed position in space. Instead of theoretically continuous flows of SD, DES uses discrete passive entities that are processed by the active entities. Products that are processed by humans and machines at an assembly line is a typical example. When implementing DES in ABM you would simply instantiate each active entity as an agent with behaviors and each passive entity as an agent without behaviors. However, in ABM the passive entities could also be made active. For example, when modelling a queue of people, an ABM would enable you to give some of them behavior like walking away, talking to each other, or starting a row.

The idea of combining SD, DES, and ABM is so logical that the firm Anylogic markets its product as a tool that can use all three paradigms [286] but as we showed this is possible in any ABM.

6.3. Create Quantified Narratives. Agent-based models make it easy to add geographic layers, agents, behaviors, and variables. The advantage is that we can avoid the falsified simplifications of equilibrium theory but it also creates a confusing amount of freedom. We think the solution is the use of clear quantified narratives. Narratives can contain a rich tapestry of variables in a way that is meaningful and memorable to the human brain, and transition management already has a tendency to tell stories [146]. We propose combining the narrative approaches used in the famous Shell scenario studies [343] with the quantitative transition management models as described in this paper that clarify, validate, and quantify the narrative.

Another reason to use the quantified narrative approach is that it enables us to explain what the model is doing to experts and policy-makers. That is needed because the integral and multidisciplinary models that we have to create in order to find the best pathways and interventions towards renewable energy require a lot of expertise. Fortunately, ABMs represent reality in a way that is easy to understand for domain experts and with quantified narratives it is also possible to explain what the model is doing. We think every ABM should be checked by experts to get face validity. This might not seem like a rigorous test but as we have seen with the example of the IEA prediction of solar power this is a test that many of the most influential models cannot pass.

7. Outlook

The transition from fossil fuels to renewable energy is a complex global challenge. Technologies like off-shore solar, airborne wind, battery storage, smart grids, and shared self-driving electric vehicles are developing quickly. They are driven by R&D, investors with foresight, users that function as early adopters, and learning curves that can become self-sustaining. Transition theory also draws attention to the ways in which innovators in niches compete with incumbent regimes and how this can lead to entirely different system configurations.

Such developments might lead to a range of very different pathways. Global warming might accelerate and this might either unite or divide humanity. Realistic carbon pricing or the price reduction of renewables might lead to stranded fossil assets and an early collapse of the fossil system. The energy system could become local and decentralized with users becoming producers that share energy and storage on demand. Mobility might become a service that replaces fossil cars by small, shared, self-driving vehicles. Developing countries without existing infrastructure might leapfrog developed countries and so on.

The modelling approach described in this paper should—at least in theory—be able to capture these dynamics in quantitative models. The agent-based modelling paradigm could also facilitate comodelling, where multidisciplinary teams of experts work together using a suite of new or existing models. Combined with object oriented programming techniques—that already allow us to collaborate on millions of lines of code—we might be able to create powerful models that uncover pathways towards a sustainable energy future that is both more realistic and more interesting than what is currently offered. We hope this paper will inspire some students, teachers, researchers, policy-makers, and entrepreneurs to do just that.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

Auke Hoekstra gratefully acknowledges the sponsoring for this research from the ElaadNL foundation.

References


