Applied Energy 305 (2022) 117949

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Spatio-temporal modelling of solar photovoltaic adoption: An integrated neural networks and agent-based modelling approach

Ali Alderete Peralta, Nazmiye Balta-Ozkan *, Philip Longhurst

School of Water, Energy and Environment, Cranfield University, Cranfield, Bedfordshire MK43 0AL, United Kingdom

HIGHLIGHTS

- Integrated agent-based and artificial neural networks model for solar PV adoption.
- The agents' adoption process is characterised using past experiences.
- Spatio-temporally explicit estimations for up to five months with accuracy of 80%.
- Integrated ABM and ANN model has adaptive capabilities over other diffusion models.
- Income, electricity usage and average household size variables yield best results.

ARTICLE INFO

Keywords:
Agent decision-making
Energy transition modelling
Technology diffusion modelling
Complex system modelling
Solar PV adoption

ABSTRACT

This paper investigates the spatio-temporal patterns of solar photovoltaic (PV) adoption, solving the ongoing need to inform the management of the distribution networks with spatially explicit estimations of PV adoption rates. This work addresses a key limitation of agent-based models (ABMs) that use rule or equation-based decision-making. It achieves this by adopting an aggregated definition of the agents using artificial neural networks (ANN) as the criteria for decision-making. This novel approachdraws from both ABM and Spatial Regression methods. It incorporates spatial and temporal dependencies as well as social dynamics that drive the adoption of PVs. Consequently, the model yields a more realistic characterisation of decision-making whilst reflecting individual behaviours for each location following the real-world layout. The model utilises the ANN's approximation capabilities to generate knowledge from historical PV data, as well as adapt to changes in data trends. First, an autoregressive model is developed. This is then extended to capture the population heterogeneity by introducing socioeconomic variables into the agent's decision-making. Both models are empirically validated and benchmarked against the Bass Model.

Results suggest that the model can account for the spatio-temporal and social dynamics that drive the adoption process and that the ABM and ANN integrated model has superior adaptive capabilities to the Bass model. The proposed model can estimate spatio-temporally explicit forecasts for up to five months with an accuracy of 80%. In line with the literature, results suggest that income, electricity consumption and the average household size variables yield the best results.

1. Introduction

A growing number of authors point out that adoption patterns modelled for domestic solar photovoltaic (PV) panels present spatial regularities [1–3]. High geographical concentrations of these inaccuracies have potential to cause problems on low voltage lines by creating reverse flows and diminishing the predictability of load, voltage and demand flows [1,4]. This modelling process shapes the evolution and characteristics of the energy system [5], as network reinforcements are required to accommodate the extra PV generation [6,7] as well as control systems to ensure stability in the voltage [8,9].

Because of the need to predict where these technologies will appear and at what pace they will evolve, the development of tools and methods that account for spatial regularities is relevant for management of distribution networks [1,10]. Yet, the diffusion of PVs is highly uncertain as their adoption is driven by subjective factors such as perceived

* Corresponding author.
E-mail address: n.ozkan@cranfield.ac.uk (N. Balta-Ozkan).

https://doi.org/10.1016/j.apenergy.2021.117949
Received 21 April 2021; Received in revised form 21 September 2021; Accepted 26 September 2021
Available online 6 October 2021

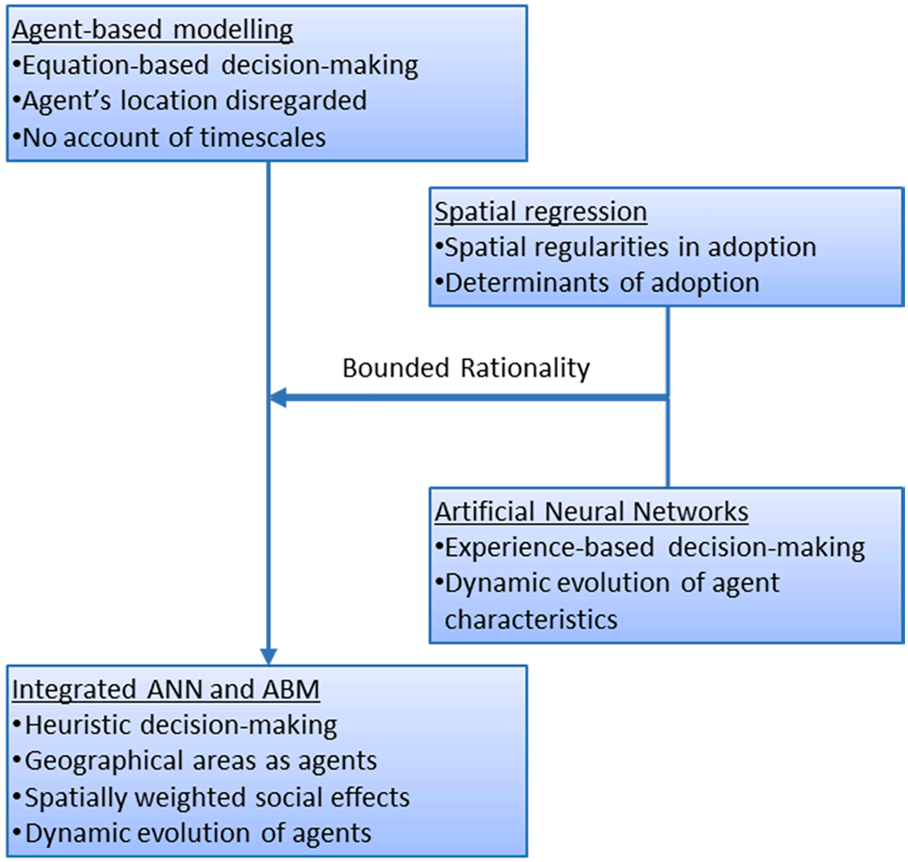

Fig. 1. Conceptual framework to develop an integrated AB and ANN model.

affordability [11], social influences from other individuals and one's neighbourhoods [1,12], or personal beliefs [13]. The adoption process is also affected by objective factors like income [2,14,15], energy costs [3,14], or available policies [16]. Current approaches to analyse the adoption of PVs fall into two broad categories. Agent-based modelling (ABM) which characterises the individual decision-making of whether or not to adopt such technologies as a result of the social interaction of the agents [17–21]. Alternatively, spatial regression[1] (SR) methods seeks to understand the effect of different factors that drive the adoption process, whilst considering the associated spatial regularities in the adoption patterns [1–3,15,22,23].

Whilst the ABM approach has been widely applied to the diffusion of innovation [24–26] and more specifically to PV adoption [27–32], this approach presents a some limitations. Firstly as ABMs focus on system-level outcomes, they don't consider the spatial location of the agents except for [28,29]. Nevertheless, these authors note that when integrating a measure of spatial regularities to the ABMs models tend to perform significantly worse [28], i.e. this phenomenon is not analysed for solar PVs. Secondly, the decision-making process in ABMs commonly assumes that agents have access to perfect market information and can evaluate the benefit of their decisions [17,33–35]. It is these challenges associated with undertaking such complicated calculations that led Noori and Tatari [36] to propose considering alternative and more realistic behavioural models by including other characteristics of human decision-making, namely artificial neural networks (ANN). ANN recognises that decision-making is driven by experience and perception rather than evaluations, as informed by the bounded rationality theory [28,37–40]. While a few studies present how ANNs can address the limitations of the decision-making process in ABMs, these are either theoretical [41] or don't fully utilise their capabilities [42]. Thirdly, as ABMs don't capture actual timescales they present limitations in characterising the dynamic evolution of agent properties as the preference's parameters are fixed [25,37].

The main novelty of this study lies in integrating ANN and ABM approaches to predict future spatially explicit domestic PV adoption patterns for the UK. This is relevant for the network operators as PVs have the potential to diminish predictability of loads, voltage and demand flows [6,43,44]. Therefore, this paper addresses the following research question:

It is possible to explicitly characterise the spatio-temporal dynamics of the decision-making towards PVs, whilst including the social dynamics?

The analysis draws from three strands of research, ABM, SR and ANN (Fig. 1) and follows bounded rationality theory to attempt to answer this question, whilst addressing some of the ABM's limitations. An integrated model is developed, implemented and validated using Birmingham as a case study. In this context, the bounded rationality theory argues that a group of individuals with similar socioeconomic characteristics and common interests can be characterised as singular decision-making units [38,45]. Using actual PV installation data from 2011, the model is used to predict domestic PV uptake at the monthly level for postcode districts in the city of Birmingham.

In the integrated ABM and ANN model, the autonomous decision-making of the agents, i.e. geographical areas, is informed by their own past decisions as well as the choices of their neighbouring agents. The results highlight three features (i) that the ABM and ANN integrated model has superior adaptive capabilities over other diffusion models, (ii) the model yields spatio-temporally explicit forecasts of up to five months with an accuracy of 80%, and (iii) that best results are achieved when including the income, electricity consumption and average household size variables.

[1] The spatial regression will be referred as SR from now on.

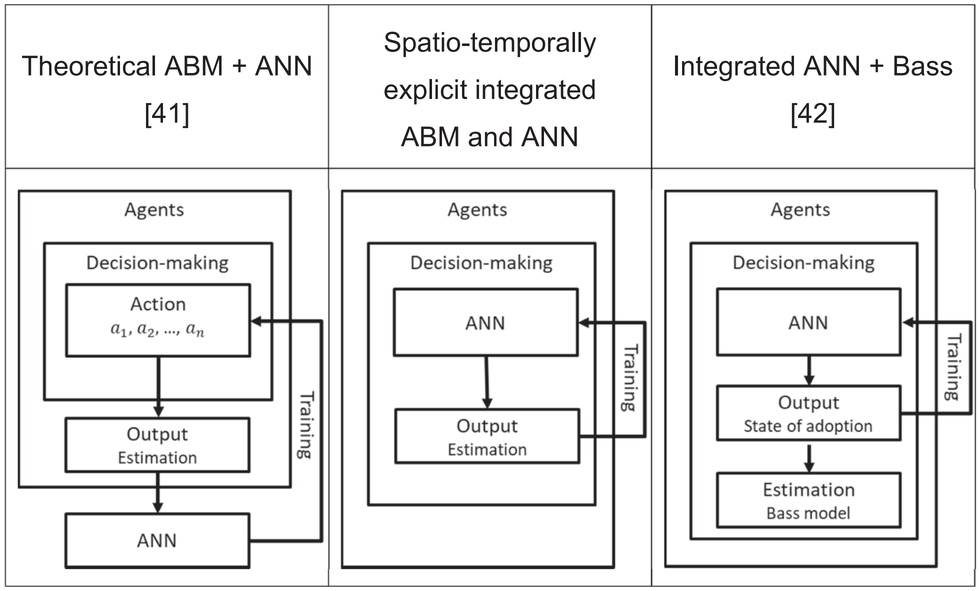

Fig. 2. Comparison of implementation of ABM and ANN integrated models.

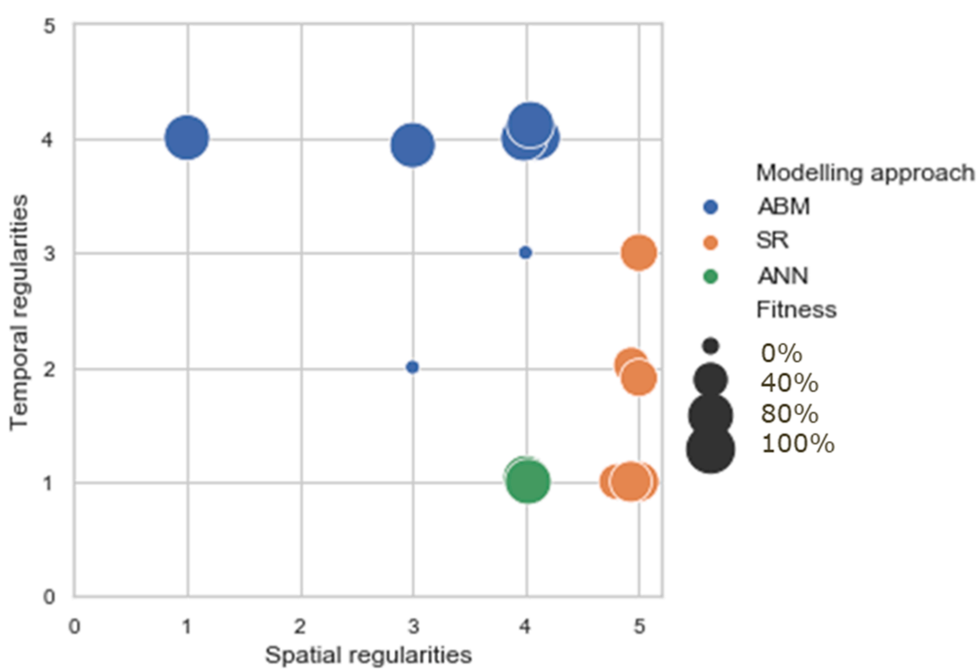

Fig. 3. Spatio-temporal assessment of ABM, SR and ANN models on PV adoption (analysed studies are: ABM - [29,32,50–54]; SR - [1,2,22,55–57]; ANN - [42,60]).

In the paper as follows: Section 2 presents an overview of the solar PV ABM models characterising the spatio-temporal dynamics of the PV adoption available in the literature. Section 3 describes the methodology to integrate ABM and ANN. Section 4 presents the results from the model simulations, which are then discussed in Section 5 with the conclusions presented in Section 6.

2. Overview of domestic solar PV modelling

ABM is one of the most commonly applied techniques used to analyse social-technical systems as it allows modelling of the entire system by characterising the role of individual actors that comprise the system [25,32]. ABM simulates the interaction of these actors through physical and social networks which cumulatively result in system-level outcomes [16,17,32]. Nevertheless, ABM faces limitations because the decision-making process assumes that agents decide by assessing financial benefits from a set of alternatives, and that social utility is based on personal and social norms. However, social utility is subjective and consumers rarely possess perfect information of the whole range of alternatives to make economic comparisons. To address these uncertainties, some authors such as Al-Alawi & Bradley [46] have run sensitivity analyses on market conditions.

Additionally, some scholars have pointed to the need for alternative decision-making methods that consider elements from human cognition [36]. Kang and Choi [41] propose an integrated theoretical ABM and ANN model as an alternative to the decision-making process. The ANN model emulates how the human brain generates information by associating experience, i.e. historical data, and the associated decision, the output. The experience-based knowledge generated reduces uncertainty as the model can adapt to new conditions that are not available during

Table 1
The criteria to assess the strengths and weaknesses of the reviewed studies.

Score	Description
1	The model disregards this modelling element.
2	The authors have implemented a proxy for this element (i.e. replicating the same analysis for different periods or locations)
3	The model can characterise this element explicitly but uses simulated or semi-empirical data to implement it.
4	The model can characterise this element explicitly but disregards its regularities (i.e. explicit location and spatial dependence).
5	The model can characterise this element explicitly and its regularities.

the training. Kang and Choi's [41] model optimises the combination of each individual's decision which is chosen from a set of possible decisions within simulated scenarios. This approach uses the ANN to optimise strategies that agents execute, i.e. evaluating the global fitness of the model to solve a specific problem. However, this decision-making is not completed by the ANN but chosen from a list of possible actions. Then, through the training process the ANN learns which combination of actions produces the best model fitness and communicates that to the agents. In this sense, the ANN interacts with the agents in a hierarchical way which is a contradiction to the ABM philosophy as agents are not autonomous [32] and the system cannot be controlled in its entirety [17].

Fig. 2 shows the proposed model (middle) using ANN for both the decision-making and estimation of the adoption rates. The ANN directly substitutes the rule-based adoption criteria by generating knowledge from the time-series and at the same time estimating future rates of adoption. As informed by bounded rationality [28,37,38], agents rely on experience and perception more than from complicated calculations, i.e. equation-based decision-making. This way we show how experienced-based ANNs' approach to knowledge can be applied to the decision-making of the ABMs whilst ensuring the agents' autonomy [47–49]. Moreover, this is expected to provide a more realistic characterisation of the adoption process, whilst providing an explicit time horizon for the decision-making (see Fig. 2).

Here we consider those ABM, SR and ANN models that have been developed to incorporate spatio-temporal patterns of PV adoption. In turn, the extent to which these models characterise the spatial and temporal dynamics of the adoption process and the corresponding models' fitness (if provided as the size of each bubble[2]) are considered. The assessment of related ABM [29,32,50–54], SR [1,2,22,55–57] and ANN [42,58] studies used to capture spatial or temporal aspects of solar PV adoption is displayed on a 5-point scale in Fig. 3. The value of 1 is used to show that either the spatial or temporal element is disregarded while a ranking of 5 reveals that this element and its regularities are characterised explicitly (see Table 1 for details). For instance, Schwarz and Ernst's ABM model [51] uses a previous survey to simulate the location of each agent, but does not consider spatial dependence; yielding a score of 3. The authors then use explicit periods for the simulations. Here, temporal dependence is not included in the adoption process, resulting in a score of 4 for temporal dependence. In general, the SR tends to oversee the time variable and thus have marginal fitness. On the other hand, the ABM approaches have limitations to represent the heterogeneity of the populations, mainly because of data availability, as mentioned in [53,59]). This also incurs limitations in representing the actual spatial distribution of the agents. However, these ABM studies can incorporate time as a variable and thus they exhibit greater fitness than the SR. An exception to this is the ABMs without a defined time horizon. These are more explorative studies than actual forecasting methods, therefore lack a fitness indicator. Furthermore, these ABMs do not consider the evolution of the agents' characteristics nor spatial dependency. Furthermore, it is noted that when integrating a measure of spatial regularities to the ABMs, models tend to perform significantly worse [28]. Despite these two models informing the potential of ANN to achieve integration of ABM and ANN [41,42], decision-making within these studies is still not fully characterised by the ANN.

The adoption of PVs is studied via ABM in a German context by Krebs [54] where an empirical model is used to characterise interactions between households. The author classifies the agent's behaviour based on socioeconomic classifications and simulates scenarios under different assumptions of ecological utility, i.e. ecological concern. This ABM characterises agents using continuous uniform distribution based on empirical data and type of lifestyle segmentations; both taken from previous surveys. The SR analyses the effect of socioeconomic variables on the adoption of PVs, whilst considering spatial regularities in these adoption patterns. Some of the recent studies include a study in the UK where Balta-Ozkanet al. [1] develop a spatial econometric method to predict spatial characteristics of the factors that affect the diffusion of PV. The authors point out the relevance of locally produced insights for policy making and make use of socioeconomic data defined in close geographical detail. For completeness, two ANN models that combine the use of cellular automata and the Bass model are also included. These models identify and classify the saturation of PV in each area using historical data and then estimate future levels of adoption using an S-curve which is discussed in Section 3.3 in more detail.

In SR studies, the cumulative number of PVs is used as the dependent variable. Most authors apply the total number of PVs, except for [2,10,57], who acknowledge the disparity in locations' size and population density. Langheim [2] normalises the number of PV installations in Southern California by considering the number of single-family buildings. Here the market size at each location is incorporated in the model. Schaffer and Brun [57] acknowledge the variation in size of locations by calculating the total PV systems divided by the location's area. The authors here use selected socioeconomic variables for the German population including solar radiation, house density, home-ownership ratio, Gross Regional Product per capita. In Section 3.2, the spatial and temporal resolution and socioeconomic variable selection are discussed[3].

Here the authors argue that the characterisation of agents and their decision-making in ABM models is a focus where key contributions can be made. The ABM models in current research characterise agents as individuals [27–29,32] or households [61,62], considering only the consumer to consumer interactions, and disregarding the merchandising activity and the seller-consumer interaction. Modellers commonly use georeferenced data (and geographical information systems) to create a world-like layout and locate the agents. The number of agents varies from a small proportion of the population to entire populations. For instance, Adepetu and Keshav [32] create a semi-empirical population that is assigned with real socioeconomic characteristics of 100 households, yet it is not clear whether the sample is statistically significant for the total population in Ontario. Instead, Rai and Robinson [27], Robinson and Rai [28], and Robinson et al. [29] model the total households from Austin (Texas, US), using data from the census. They then differentiate adopters from non-adopters, applying the findings from a previous survey carried out for PV adopters.

An alternative approach combines current survey data with other methods or software to create a virtual population. Survey data has been used to characterise significant samples of the population, which is then simulated as many times as needed to capture the actual population size [14,50,52]. Similarly, Cui et al. [61] generate a virtual population of households based on actual aggregated data. By applying the *copula-based household synthesiser*[4], they generate an individual virtual

[2] In cases the fitness isn't provided explicitly, it's set to 40% by default.

[3] For a recent review of SR studies, the interested readers may wish to see [90].

[4] This method simulates households using local distributions (i.e. the statistics of each census block), which allows to keep empirical correlations.

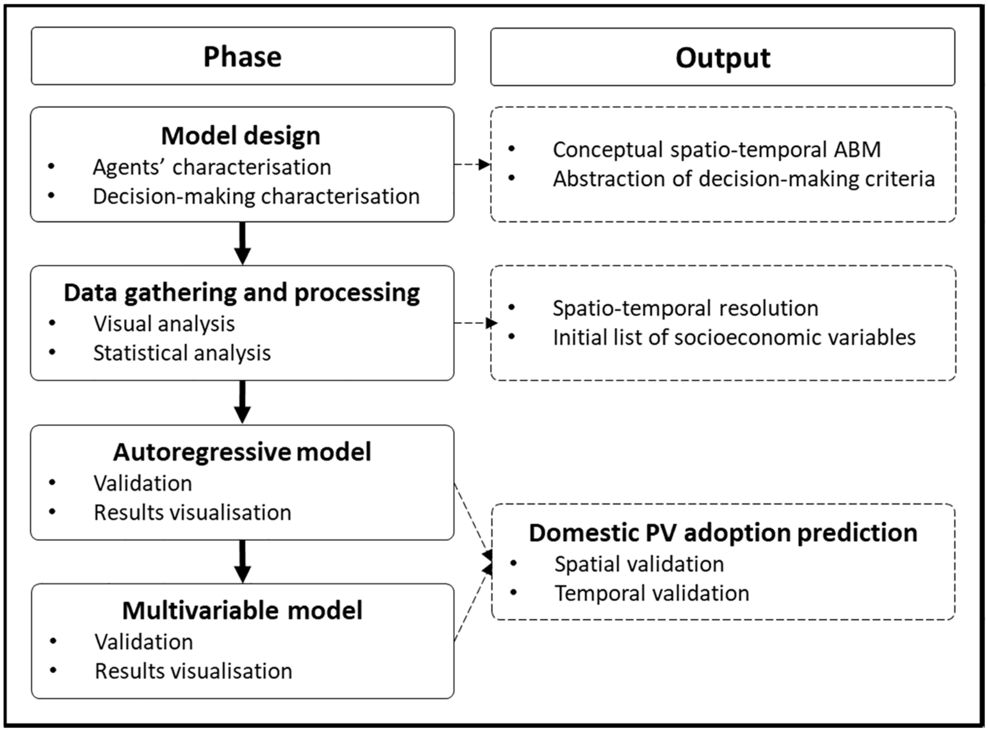

Fig. 4. 2-phase methodology for the implementation of an integrated spatio-temporally explicit ABM and ANN model.

household population with similar characteristics (with intra-group variance), resulting in the simulation of 190,965 agents. Eppstein et al. [62] create a full virtual population of households. First, they generate a virtual distribution of *income*, considering five hypothetical cities. Drawing from previous experimentations showing no significant differences between simulations with 1,000 or 10,000 agents, the authors use 1,000 in the interest of computational efficiency. These agents are created using the *turning bands method*. The location of the agents is fully related to the agents' characterisation, yet empirical or semi-empirical characterisation may imply a simulated location. Hence, the actual location may differ from the simulation and may involve a loss in data accuracy [14]. Although 'the agents' location needs to be accurate, it is a challenging task to obtain the personal characteristics of the entire population [61].

Commonly ABMs use rational choice principles where one assumes that agents have access to perfect market information and can evaluate the benefit of their decisions. However, this is rather limiting as some of the drivers such as peer-effect or personal beliefs have subjective value, and individuals rarely possess perfect market information [17,33–35]. Authors then use a utility or social threshold to characterise their decision-making process [62,63]. Such utility thresholds consider exogenous and objective elements to the agents, e.g. electricity prices or government subsidies and this is usually associated with the financial benefits of adopting a certain technology. On the other hand, the social threshold reflects subjective elements such as the agents' personal beliefs, values, and therefore reflects how adopting a certain technology satisfies the interpersonal preferences of the individual.

Agents' decision-making process are commonly defined by a single criterion [14,16,32,61,63]. Yet some studies employ a multi-step process, including Eppstein et al. [62]. These characterise agents with an equation to calculate the relative cost of, for example buying a new vehicle amongst a set of options. They consider the purchase cost, financing, fuel and electricity cost. The agent then compares this cost with net annual income. If the sum of these costs exceeds 20% of the annual income, this is considered unaffordable. However, if the vehicle is affordable the agent then considers the social benefits. The agent then calculates the perceived social benefit of each vehicle; if these benefits exceed the social threshold, the agent adopts that specific option. Others such as Robinson et al. [28,29] model the decision-making process using both social and utility thresholds. Financial benefits are calculated based on the financial payback, considering the energy produced by the solar panel, the price per kWh generated, and the government subsidies. In this context, social threshold is related to the attitudinal variables which change according to the social network effect.

Regardless of the type of threshold used to characterise the decision-making process most ABM studies implement the interaction among agents. These interactions can be described in two ways, the peer-effect and the social norms. The former refers to the direct influence of the social network in the decisions (perception). The latter comprises the impact of collective behaviour on the agent's decision-making. The agent's social network may be characterised using different approaches, for instance, defining a spatial neighbourhood or choosing between those agents with similar characteristics. Eppstein et al. [62] define the agent's social network as the k nearest neighbourhood (starting with $k = 2$) and generate random social networks. Similarly, Adepetu, Keshav and Arya [14] and Schwarz and Ernst [51] define spatial proximity and connect the agents with similar socioeconomic characteristics. Robinson and Rai [28] combine the distance-based and agents similarity criteria with a random connection so that agents are connected with other agents anywhere in the area. Exceptionally, Cui et al. [61] do not provide detail on the agents' social network structure. Alternatively, the social norm has been modelled by considering the adoption rate of the total population as a proxy for the social effect [16,32,63]. Similarly, Eppstein et al. [62] use the media coverage in an area as an indication of the environmental awareness which increases over time, whilst Robinson and Rai [28] implement the perception of the technology as a measurement of social awareness of the social norm.

Regarding the ANN studies, Zhao et al. [42] and Liu et al. [58] use the ANN only as decision-making, whilst the estimation of adoption rates is done using the Bass model. In this study, the neural network model is trained to memorise the socioeconomic characteristics of the areas with PVs, and then identify when a location reaches these

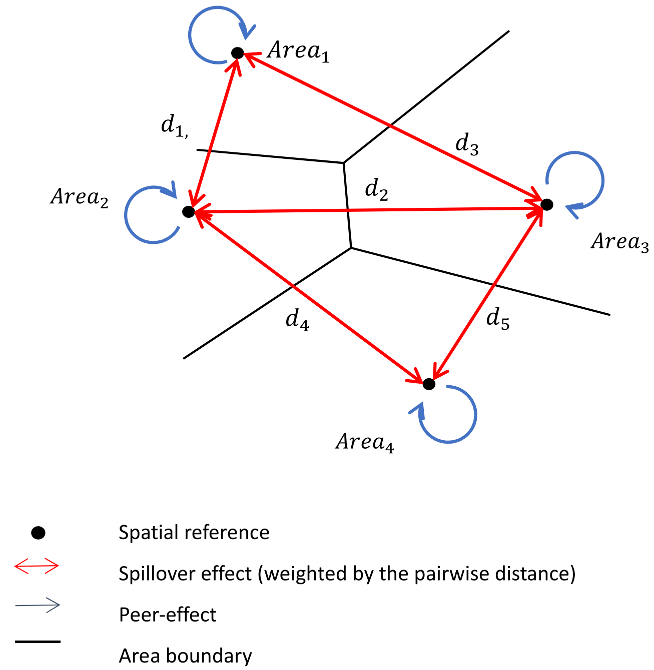

Fig. 5. Conceptual model of a spatio-temporal explicit ABM.

characteristics. The areas in this model are uniformly defined grid cells. Real-world layout and impacts of areas on each other are disregarded. The forecasting process calculates the total amount of PV in each region using an S-curve, which shows the historical growth of the PV in the entire population. The assignment of local values from regional characteristics does not allow the model to consider local regularities.

The social dynamics that drive the adoption process also present spatial regularities. Namely, these social dynamics include the peer-effect and social norms and capturing the influence that one individual has on another or the influence of the overall society's trend, respectively. Despite the limitations of the ABM to implement the spatial dependence, some authors have defined the agent's social network using the distance between agents as criteria to connect two agents [51,62]. However, whether the distance between agents reinforces or decreases the social effects has not yet been tested.

In contract, the SR allows explicit characterisation of the spatial layout of the diffusion process. Despite this, these approaches tend to overlook the temporal dimension of the adoption process. In spatial regression, the so-called spillover effect captures the information flow between or within specific locations [1,12,22]. This effect is implemented in different means, for instance, Balta-Ozkan et al. [1] characterise this as the influence of one area over the adjacent locations. Conversely, Richter [12] defines this as the transmission of information between individuals in the same area. Like the ABM, areas are influenced by those in their social network, despite this social network being based on geographical elements. Spatial regression models identify this relationship in two ways: adjacency and distance. In the former, this influence has the same magnitude for all areas that share a border, whilst disregarding the influence of areas that are not adjacent [64]. The distance-based approach disregards the adjacency and assumes that the social influence decreases with the distance such that the influence of one area over another will have a greater impact the closer they are. Thus, this research argues that the spill-over effect as a representation of spatial dependence can be homologous to the social influence of the ABMs.

Step-by-step development of an ABM model that implements the spatio-temporal and social dynamics of the adoption process is explained in the following section.

3. Methodology

Here the authors propose an integrated agent-based modelling and artificial neural networks to characterise the spatio-temporal adoption patterns of PVs. As illustrated in Fig. 4, the analysis is performed in four stages. Because of the complexity of constructing this new approach, the first step is to design an autoregressive model for the PV adoption process. This stage focuses on defining the characterisation of agents' and their decision-making, and showing how the model captures the different adoption factors and the general structure of the ANN. The second step includes data collection and processing, including spatial mapping and statistical exploratory data analysis. This stage defines the resolution of the data, i.e. the number of agents and time-steps in the simulation (total number of time series observations). The third and fourth steps comprise the implementation of the autoregressive and multivariable models and the analysis of the results after the training of

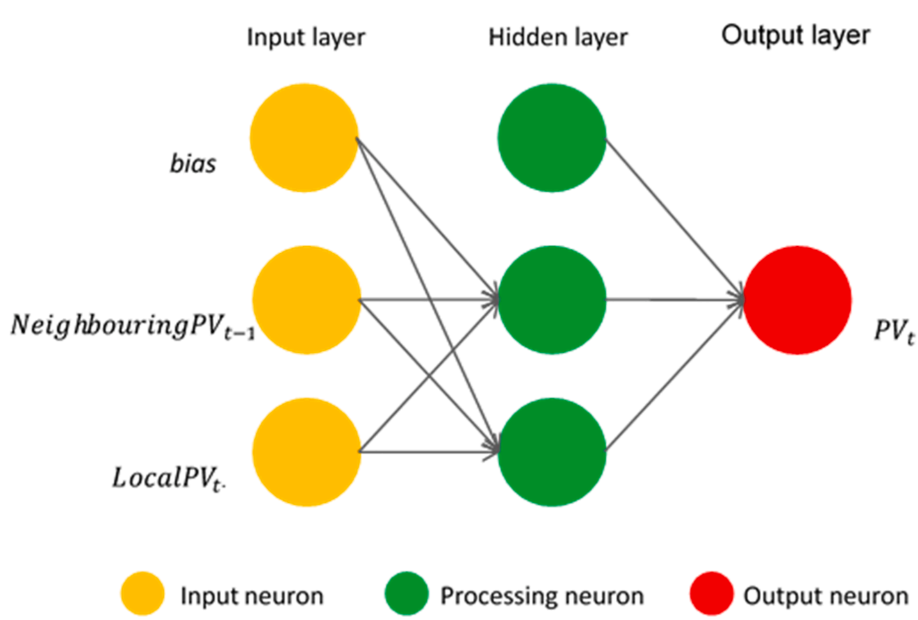

Fig. 6. Artificial neural network design for a spatio-temporal explicit ABM.

Table 2
List of independent socioeconomic variables and their resolution.

Variable	Spatial resolution*	Temporal resolution	Data points	Data source	Aggregation		Disaggregation Temporal
					Spatial	Temp.	
PV installations	LSOA	Daily	2011–15	OFGEM	X	X	
Weekly income	MSOA	Weekly	2013, 2015	ONS	X	X	
Population density	PC	Census	2001,2011	ONS			X
% Owned household	PC	Census	2001,2011	ONS			X
% Detached household	PC	Census	2001,2011	ONS			X
Electricity consumption	LSOA	Annual	2001, 2011–15	ONS	X		X
Education level	PC	Census	2010,2015	ONS			X
Average household size	PC	Census	2001,2011	ONS			X
CO_2 emissions proxy	LSOA	Annual	2001,2011	ONS	x		X

the ANN. Although both steps comprise the validation and presentation of their results, the multivariable model is only implemented once the autoregressive model has been validated.

3.1. Model design

Fig. 5 illustrates the configuration of the (i) spatial and (ii) temporal dependence, (iii) peer-effect, (iv) spill-over effect of a spatio-temporally explicit ABM and ANN approach. The agents are geographical areas ($Area_i$) that represent the cumulative decision making of the individuals living in that area [12,38]. The model accounts only for the end-users interactions and ignores any activity done by the sellers/car retailers [1,12,32,55]. Eq. (3.1) abstracts the agents' decision-making, which estimates the number of PVs considering these elements. Eq. (3.2) shows the extended characterisation which includes socioeconomic variables.

$$PV_t = f\left(LocalPV_{t-1}, \sum d^{*}NeighbouringPV_{t-1}\right). \tag{3.1}$$

$$PV_t = f\left(LocalPV_{t-1}, \sum d^{*}NeighbouringPV_{t-1}, Socioeconomic_{t-1}\right). \tag{3.2}$$

where $\boldsymbol{PV_t}$ is the total number of PVs in a specific time, $\boldsymbol{LocalPV_{t-1}}$ is an autoregressive element (peer-effect), $\boldsymbol{NeighbouringPV_{t-1}}$ is the number of PVs in the adjacent areas (neighbourhood effect), $\boldsymbol{d}$ is the distance between the locations and weights the social effects (spatial dependence), $\boldsymbol{Socioeconomic_{t-1}}$ is a set of independent variables, and the temporal dependence is denoted by lagging these inputs in ***t-1***.

3.1.1. Agents' decision making

Decision-making by the agent (Eqs. (3.1) and (3.2)) is characterised by an artificial neural network. ANNs are like linear (and non-linear) least squares regression, in the sense that both attempt to minimize the sum of squared errors. In this case, the fitting takes place through a training process, which consists of presenting pairs of inputs and outputs, calculating the error of estimation, and subsequently adjusting the ANNs' weights. Fig. 6 shows the structure of the ANN used as the agents' decision-making, linear and sequential structure, with three layers of neurons, the input, the hidden and the output layers. A linear structure with a single hidden layer has been proven to be sufficient to approach any nonlinear function [65] and used to forecast the PV generation [58,66].

The input layer (yellow) includes a node for each of the social effects for t-1; the hidden layer (green) comprises one neuron for each of the inputs, and the output (red) layer contains a single neuron to produce a single output. The multivariable characterisation increases the number of input neurons, which depends on the number of socioeconomic variables selected. The neurons use an activation function, which is similar to the biological activation threshold [47,67,68]. No calculations are made at the input layer; then, because the model accounts for the decision of whether to adopt or not (a binary decision), the hidden and output neurons the model implements the sigmoid function, and the learning rate is 0.1 [47,58,69]. The synaptic weights that connect the layers are randomly initialized with values between [0–1]**.** Additionally, the model includes bias nodes in the input and hidden layers [47,58,69]. These are nodes with a constant value of (1) and can be interpreted as the β_0 in a Linear Regression; which is a constant term reflecting the intercept of the function.

The training uses the Backpropagation algorithm, which is a two-phase process. During the forward phase, inputs from the training subset are passed through the neural network to determine the output. Once an output has been estimated, an error of estimation is calculated against the expected value. Then, this error is propagated through the network in the backward phase, by adjusting the weights and minimizing the error of estimation. Once the whole training subset has been presented, the model validates the ANN's estimations against the validation subset. If the ANN does not meet the stopping criteria, the process is repeated, presenting the whole training subset again. Commonly, the stopping criteria are related to the level of error; in other words, if the ANN can estimate the actual data at a certain accuracy level then the training is over. For a detailed mathematical description of the backpropagation see [47,69,70].

3.2. Data gathering: Spatial and temporal resolution and variable selection

The analysis builds upon Balta-Ozkan et al.'s spatial-econometric model [1] for the diffusion of PVs in the UK. To identify the list of socioeconomic variables that affect the adoption process [71], the econometric stepwise method is used to assess how each variable predicts the adoption rate. Table 2 shows the characteristics and data sources of the variables included Balta-Ozkan et al.'s model [1]. PV data is available in the Feed-in Tariff Installation Report published by the Office of Gas and Electricity Markets (Ofgem) quarterly, which contains the registration date (dd/mm/yyyy) of domestic PVs. Socioeconomic data is published by the Office of National Statistics, containing data such as income and homeownership which are available for 2001 and 2011. Appendix A summarises the statistics of the potential explanatory variables, presenting min, max and range of values. The boxplots for the variables show that as more PCs are included, the statistics fluctuate. It is expected that the ANN adapts to a wide variety of data behaviour, improving the confidence in the model.

Because the spatio-temporal resolution of some of these variables' is not in the required resolution, the data is either aggregated or disaggregated to produce monthly observations. Table 2 shows which treatment was given to the variables to meet the spatial resolution requirements. The variables at Lower Layer Super Output Area (LSOA) or Medium Layer Super Output Area (MSOA) level were aggregated to PC level using Office of National Statistics reference lookup tables[5]. The

[5] Office of National Statistics Source (accessed on 16/09/2020): https://ons.maps.arcgis.com/home/item.html?id=ef72efd6adf64b11a2228f7b3e95deea

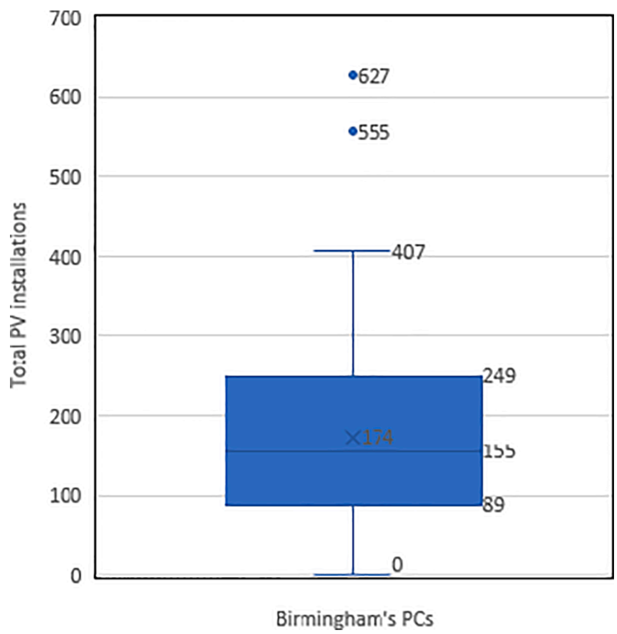

Fig. 7. Boxplot of the total PV installations in Birmingham.

only exception is the solar irradiation data because the changes in solar irradiation between one area to another, in this study scope, is negligible[6]. A similar process is taken for the variables with higher temporal resolutions, these are aggregated on monthly basis. On the other hand, data for the variables that are available annually or included in the census are temporally disaggregated following the UK Office of National Statistics (ONS) methodologies for temporal disaggregation [72]. Temporal disaggregation is a process that generates a time series at a higher frequency from data with a lower temporal resolution. Monthly gross domestic product (GDP) observations have been estimated from the annual time-series, by applying Fernandez' technique and using the Index of Services[7] [72]. Then, following the ONS methodology, monthly observations are estimated with the Fernandez algorithm. As this estimation is at the national level, the Index of Services is replaced by an index of house pricing which has a high spatio-temporal resolution (monthly/LAD).

3.2.1. Case study

To illustrate and validate how this approach works in a real-world scenario, this analysis uses historical PV installation data for the City of Birmingham in England. The *Feed-in Tariff Installation Report* accounts for the geo-referenced[8] database of all renewable energy installations such as anaerobic digestion, combined heat and power, hydro, wind, and solar photovoltaic. The analysis comprises data from Jan 2011 to Dec 2015, resulting in a total of 60 observations (months) for the 49 postcodes districts (PCs) in Birmingham.

Over this period the number of domestic PVs installed increased from $N = 30$ to 12,148[9]. Fig. 7 shows the range of values among the PCs, which have an average of 151 PVs. Despite most of the areas having a total number of PVs between 100 and 200, there are areas with extreme values (more than 400 and 0 PVs), resulting in a large variance in the data across the areas. It is expected that that given the individuality of the ANNs, the agents can adapt to the local variations, instead of fitting their behaviour to the overall data trend. The spatial distribution of PVs at PC resolution is shown in Fig. 8. As seen, there are a few areas with less than 50 PVs, most of which correspond to the city centre areas where the number of residential buildings is low.

3.3. Model validation criteria

The analysis assesses the model's performance in space and time using the Mean Absolute Percentage Error (MAPE) [47], which measures the overall performance over a specific time horizon (Eq. (3.3)). This is a widely used measurement of error in a variety of areas such as energy usage/output forecasting [21,58,73,74], diffusion of technology [42,60], or financial forecasting [75,76]. Because the model comprises a population of agents, their individual performance is aggregated and weighted against the population size. The Individual and population's MAPE's definition is shown in Eqs. (3.3) and (3.4).

$$MAPE_j = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{PV_t - \widehat{PV}_t}{PV_t}\right|. \tag{3.3}$$

$$Population\ \ MAPE = \frac{1}{m}\sum_{k=1}^{m} MAPE_k. \tag{3.4}$$

where $\boldsymbol{n}$ is the time series size, $\boldsymbol{PV_t}$ is the current number of PVs in the month $\boldsymbol{t}$, $\widehat{\boldsymbol{PV}}_t$ is the estimation of the number of PVs in the month $\boldsymbol{t}$, $\boldsymbol{i}$ is the specific month (among the time-series), and $\boldsymbol{j}$ is the specific postcode district. Then, for the *Population MAPE*, $\boldsymbol{m}$ is the population size, $\boldsymbol{k}$ is the specific area (among the 49 PCs). Additionally, the Bass model is used to calculate the saturation curves (s-curves) and benchmark to validate the ABM and ANN models. The Bass model is commonly used to model the adoption of innovation [46,77,78], assuming that these technologies are adopted by imitation or media-effect influence [46,79]. Al-Alawi and Bradley [46] and Heymann et al. [80] mention that the Bass model is the most extensively used approach for the adoption of innovation. Despite the academic community has extended the Bass model to capture the complexity of communication channels and trends of adoption of new products [81], the model does not consider the spatial dependence and have limited capabilities to account for the temporal dynamics of the adoption process [82]. The analysts estimate the market saturation s-curves using the same datasets as described in Appendix B; which summarises the methodology followed; the errors of estimation are calculated in the same way as those for the ANNs.

3.4. Model implementation

The simulations follow the four steps procedure shown in Fig. 9, which starts with the agents' characterisation, creating an agent for each of the PCs in the Birmingham area; the pseudocode of this procedure is presented in Appendix C. The model uses boundary files (.shp) to define the shape of each area and to assign a georeference according to the population centroid. Each of these areas (agents) makes a query to the PV installation database, loading their time-series. Using the adjacency principle, the model associates neighbouring PCs and calculate the distance between their centroids. This is then used to calculate the spill-over effect. Afterwards, each of the agents is assigned with a neural network, which is initialised using random values.

In the second step, the ANN is trained by presenting pairs of inputs and outputs to the neural network where knowledge is generated through the learning algorithm. Because the decision-making considers the actions of other agents, the algorithm must communicate each

[6] The distance between Birmingham and Edinburgh is ~394km, whilst the horizontal solar irradiance changes only by less than 10% [88]. Yet, the longitude of a cross section of Birmingham is around 30km.

[7] The Index of Services measures the quantity of output from all UK services industries, and accounts for more than three-quarters of the output approach to the measurement of Gross Domestic Product.

[8] Each registration is geo-referenced to a postcode, which can be later be associated with other spatial resolution, such as Local Authorities.

[9] Note that because of the low level of PV adoption before Jan 2011, this data was excluded from the analysis (Total PV < 3% before 2011; Average PV per PC < 4)

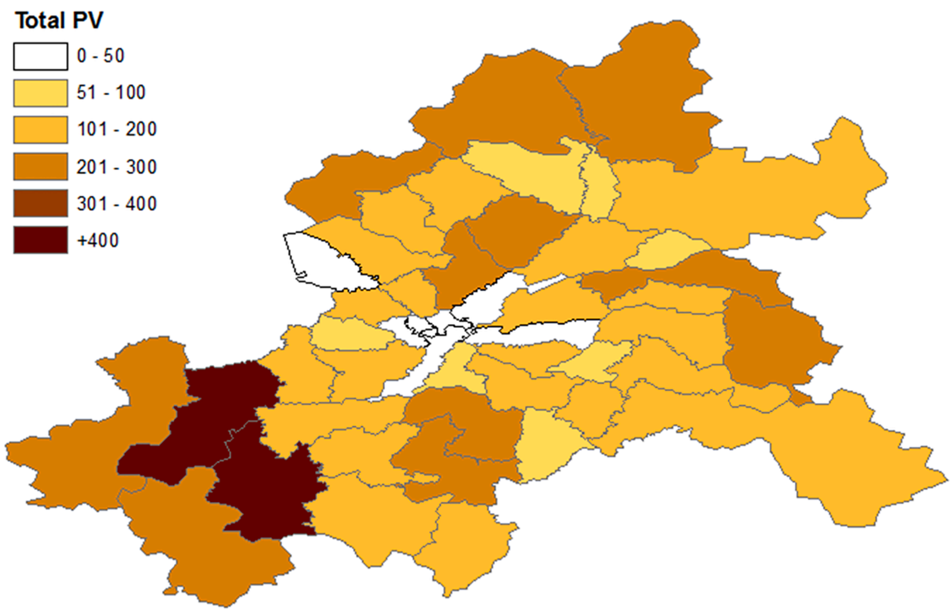

Fig. 8. Spatial distribution of the PV installations in the Birmingham postcodes (Dec 2015) (All maps presented in this thesis are self-elaborated using the data stated in each of the chapters, using the ArcMap v10.4.1 software.).

decision to the population. The information flows occur considering the social network of each agent and their distance. To reduce computation time during the training, instead of calculating the number of PVs in the surrounding areas, this number is calculated beforehand from the raw data and included in the database. Thus, the time-series for training comprises the total number of PVs in the adjacent areas which are weighted by the distance between the areas. The training validation is an internal process of this step and is independent of the overall model validation. In this case, if the ANN's error meets the stopping condition (training error < 0.001), the training stops, otherwise the training continues.

The forecast process takes place in steps three and four. First, each agent estimates their PV adoption rates, then they communicate these estimations to the agents in their social network. Because of the limited number of observations (N = 60), the time-series are divided into 90%, 5%, 5% sets for the training, validation and forecast [47,67,68]. After optimising the ANN synaptic weights, the training and the validation subsets are merged, enlarging the training set (95–5%). Given the random element of the synaptic weights, the model runs the whole simulation 100 times, then, the results report the average behaviour of the model's output.

The model was implemented using the *AnyLogic* v7.3.2 software, Java programming language for the ANN module, Python 3.5 and ArcMap v10.4.1for the statistical tests. Algorithm 1, in Appendix E, shows the pseudocode for the step-by-step main procedure for the simulations.

4. Results

In this section, we present results from the autoregressive and multivariable ABM and ANN models in comparison to the Bass model where relevant. Firstly, the temporal results are presented to visualize their behaviour over time. Then, the spatial distribution of the results is presented either at the end of the training period or specific forecast periods.

4.1. Presentation of results temporally

Fig. 10 shows the PV adoption rates estimated by the autoregressive, multivariable and Bass models, together with the actual PV values. The differences in performance of these three approaches are clear during the mid-periods characterised by large deviations. The ANNs' estimations follow the actual PV installations closely most of the time, however, the ANNs are more likely to produce extreme values at the beginning of the training because the neural network has been fed with a small proportion of the time-series data [65,67,68,83]. The results reported are the average of the 100 simulations. The randomness in the ANNs initialisation means that from the results, it is recognised that there could be a best and worst case scenario. Fig. D1 shows the average MAPE for the 100 simulations, as well as the best scenario and worst scenario. As seen, these last two follow the same patterns like the average, suggesting that even with the random element the training algorithm converges.

At the end of the series the autoregressive ABM model produces significant larger estimations; at a rate of eight times higher than the multivariable model. On the other hand, the Bass model is closer to the actual data at the beginning and at the end of the time series, when the data has no deviations. However, the Bass estimations diverge from the actual values in the middle of the series, as this model has no adaptive capabilities.

Among the three models, the multivariable ABM produces the estimations closest to the actual behaviour. These results are yielded when the **income**, **electricity consumption** and **average household size** are included. This combination of variables increases the accuracy of estimations to 95%. As expected, the addition of socioeconomic variables yield better results than the autoregressive model [47,84]. Similar to the SR, these added variables explain the variance of the dependent variable. The variables included in the multivariable model are in line with those that have been proven to drive the adoption process in SR studies. For instance, **income** has been used both to define the agents' utility or social threshold in ABM studies [14] and as a key variable determining the decision to adopt the PV technology in SR studies [1,7,85,86]. The **electricity consumption** variable reflects the concerns of households with high energy usage to become self-sufficient [1,7]. **Average household size** has a statistically significant negative impact on PV diffusion, as bigger families may have less cash flow for a PV installation [1,15,55]. Table F1 in Appendix F shows the stepwise process of variable selection, which in this case stops at the fourth step.

To visualise the adaptation of the ANNs to the disturbances during the middle of the time-series, the analysis calculates the absolute marginal change of the MAPE over time and the average of those changes,

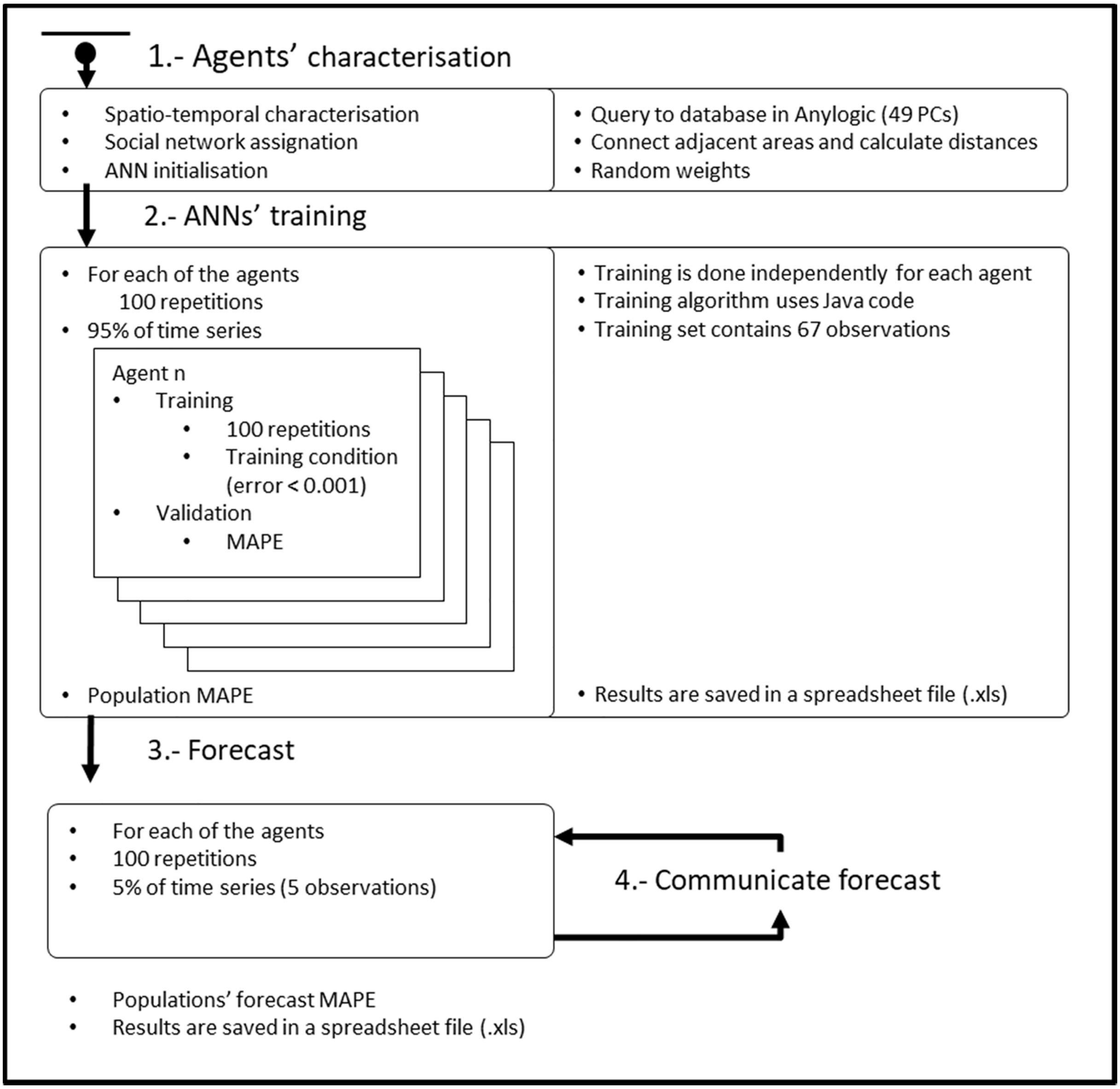

Fig. 9. A four-step process for the implementation of the model and simulations.

using Eqs. (4.1) and (4.2):

$$Absolute\ \ marginal\ \ change_i = |MAPE_i - MAPE_{i-1}|. \tag{4.1}$$

$$Average\ \ marginal\ \ change = \frac{1}{n}\sum\nolimits_{i=1}^{n} Absolute\ \ marginal\ \ change_i. \tag{4.2}$$

Fig. 11 shows the distribution over time of the marginal changes that were higher than the average. The autoregressive residuals show three major disturbances or sharp changes in the agents' behaviour. In contrast, the multivariable model presents only one major disturbance at the end of November 2011. The concentration of these changes suggests that some of the agents' behaviour is not captured either by the autoregressive or multivariable models and that there might be other factors that are driving this shift. One hypothesis could be that these peaks might be linked to the changes in Feed-in Tarrif (FiT) rates which have been shown to play a significant impact on PV adoption [12,87–89]. Moreover, these disturbances are not present in all the PCs, thus, it can be argued that the effect of the FiTs rates may present spatial regularities. This highlights the importance of considering local socioeconomic factors when designing new policies [90]. Fig. 12 shows the marginal changes in the total number of PVs for each PC from October to November 2011, one could argue that this reflects the local responses to the FiT. Building on this assumption, the spatial distribution of these marginal changes are shown in Fig. 12. As shown, the most responsive areas are those at the edges of the city, while PCs near the centre are less responsive.

4.2. Presentation of results spatially

Fig. 13 shows the distribution of the estimation errors for both ABM and Bass models, respectively. Here, the Bass model yields a more uniform distribution of the errors, with most of them between 11% and 30%. On the other hand, the ANN estimations present a more random

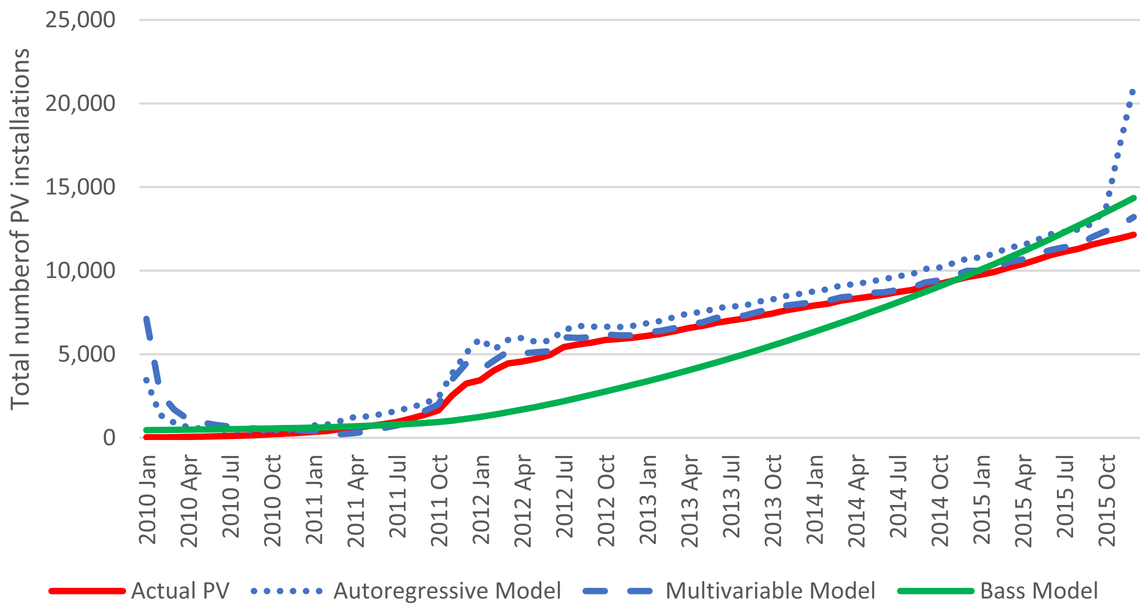

Fig. 10. Cumulative adoption rates of PVs estimated by the ABM and the Bass model vs. actual data.

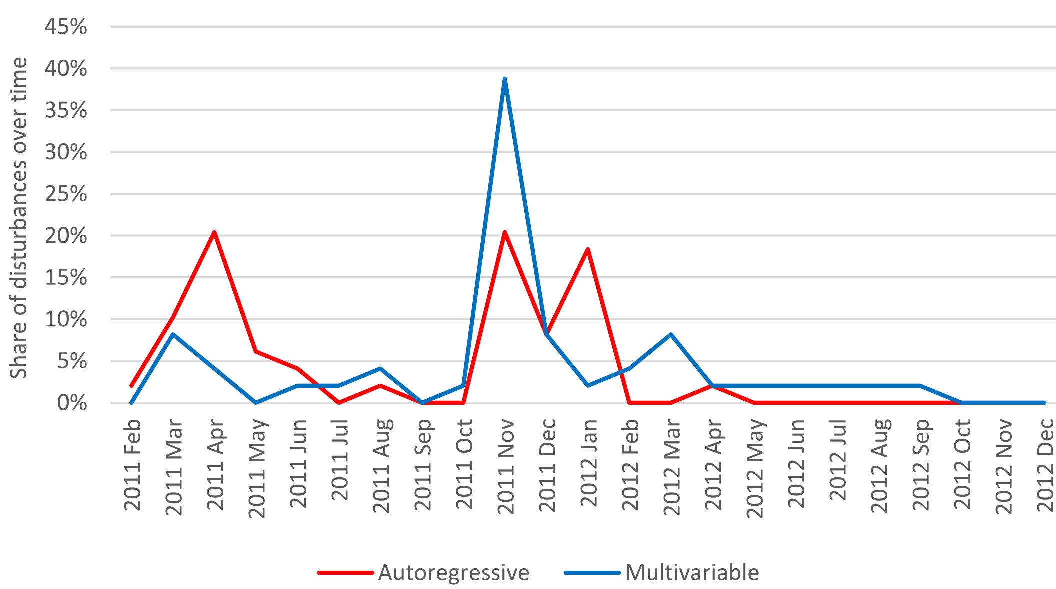

Fig. 11. The temporal patterns of the marginal changes of the estimation errors over time – autoregressive and multivariable ABM and ANN models.

spatial distribution of the estimation errors, with most of them between 1% and 10%. For the three models, the largest estimation errors are present in the city centre PCs (Fig. 13a, b, c). Since these PCs have a low density of residential buildings, they also have a low or null number of PVs. Because the MAPE calculation is sensitive to any minor change in small numbers, the relative under or overestimation in one unit will produce a larger error than in areas with a higher number of PVs. To assess the significance of the spatial patterns of the estimation errors, the results are analysed by considering the spatial regularities of the errors using the hot spot analysis [91,92].

Fig. 14 presents the maps for the hot spot analysis for the three estimations. In both cases, most of the areas present non-significant regularities except for the central PCs, meaning that the MAPE is mostly randomly distributed and that the model can capture most of the spatial dependence [92]. The clusters exhibited by the estimation errors correspond to the areas with low residential buildings, as mentioned before.

4.3. Domestic PV adoption prediction

Fig. 15 shows the five forecasted months of PV adoption across the integrated ABM and ANN and the Bass models. As seen, both ABMs

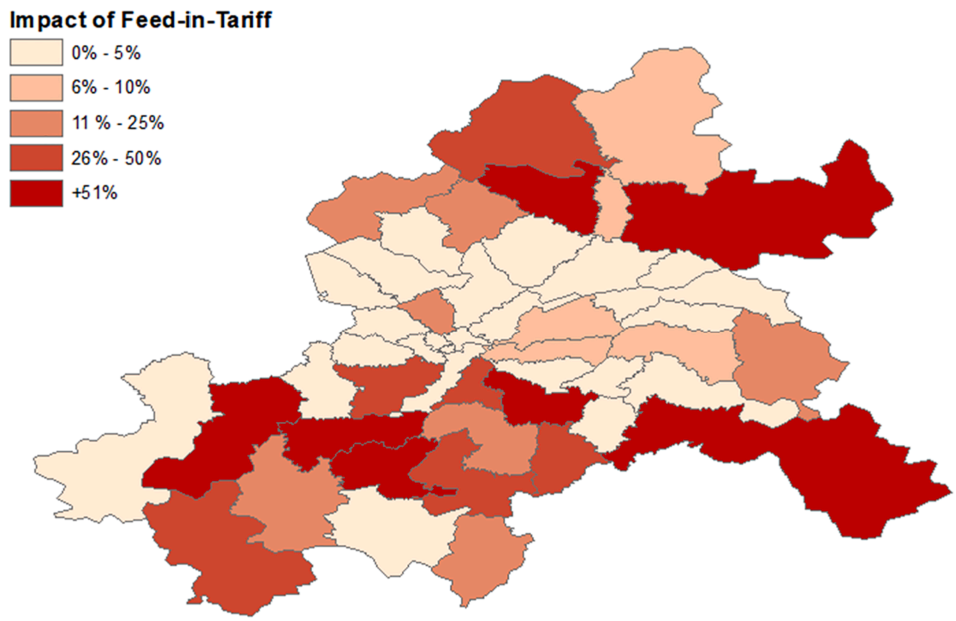

Fig. 12. Spatial distribution of the assumed impact of FiT on the adoption rates by November 2011.

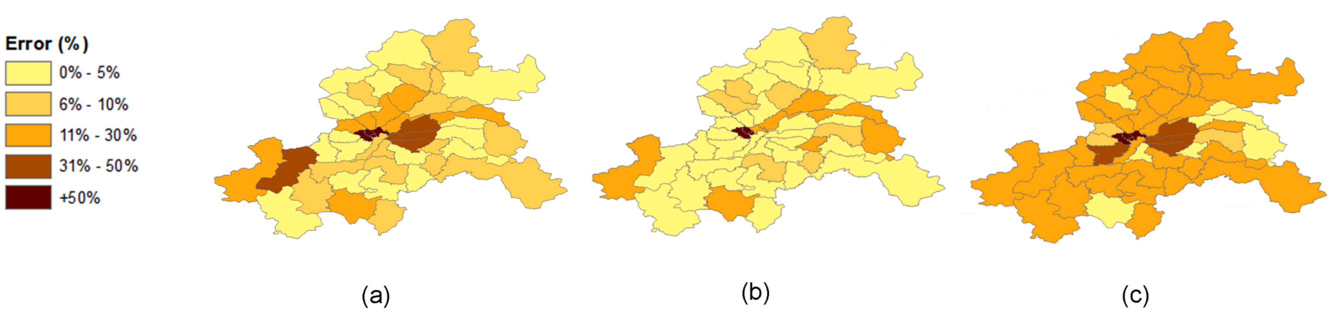

Fig. 13. Spatial distribution of the estimation's error by the end of the training - a) ANN autoregressive, b) ANN multivariable, **c)** Bass model.

Fig. 14. Hot spot analysis of the estimation's error by the end of the training – d) ANN autoregressive e) ANN multivariable, f) Bass model.

exhibit opposite behaviours: Firstly, the estimation errors of the autoregressive model quickly accumulate reaching levels higher than 70%. Because the forecast is completed dynamically after the training, each estimation considers also the estimation errors. Thus, the total number of PVs in a given period may accumulate errors from all estimations for the preceding periods. Yet, as expected, the multivariable model accumulates the errors at a significantly lower rate. This is because the multivariable model reduces the error's magnitude at the end of the training, i.e. having a 10% error at the end of the 5th period against 75% error accumulation of the autoregressive model. While the error levels of the Bass model present a more stable behaviour, they are greater than those for the multivariable one. As the Bass model disregards the temporal dependency, its estimations are independent of each other and the errors do not accumulate.

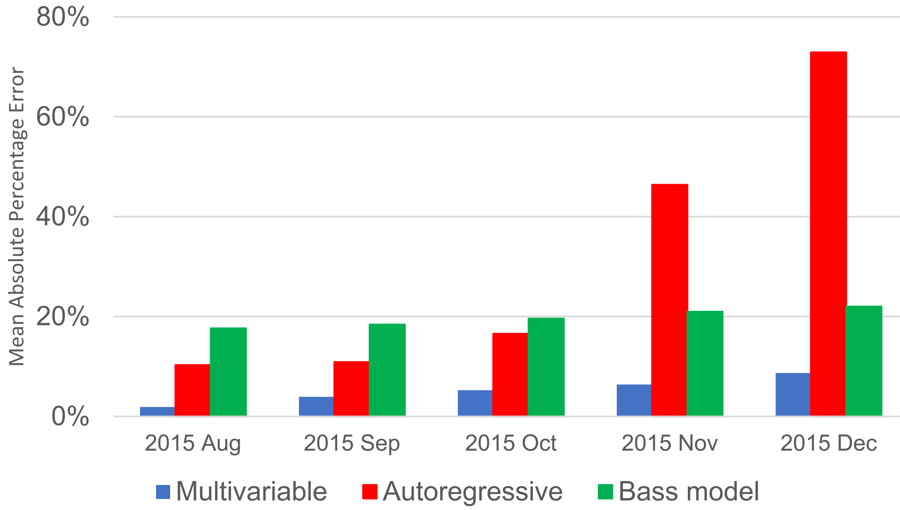

Fig. 15. Estimation error for the forecasted periods - autoregressive model vs. multivariable.

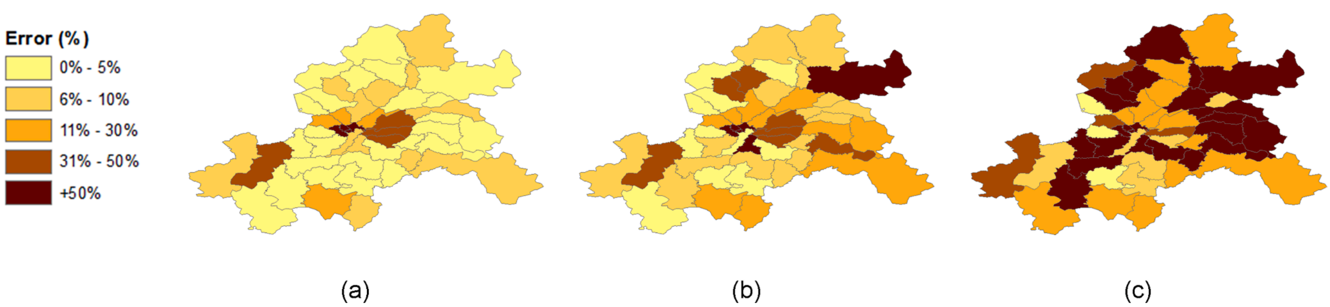

Fig. 16. Spatial distribution of the autoregressive ANN estimation's error – (a) First, (b) Third and (c) Fifth forecasted period.

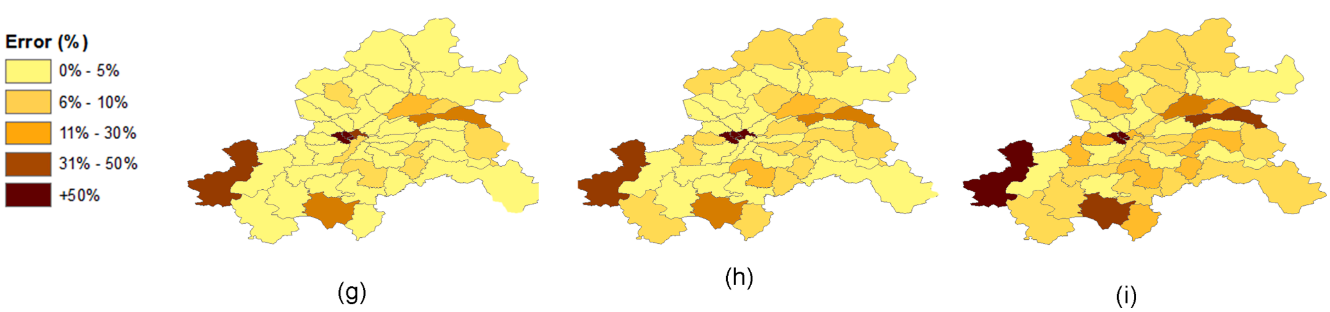

Fig. 17. Spatial distribution of the multivariable ANN estimation's error – g) First, h) Third and i) Fifth forecasted period.

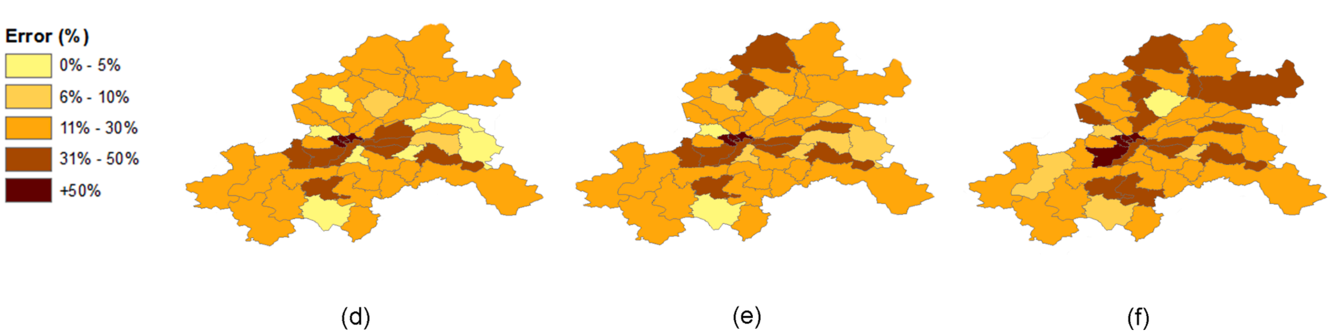

Fig. 18. Spatial distribution of the Bass model estimation's error – d) First, e) Third and f) Fifth forecasted period.

The spatial distribution of the ANN's estimations by PCs are shown in Fig. 16, Fig. 17 and Fig. 18. These follow the spatial distribution of Birmingham's PCs. Fig. 15 shows the aggregated MAPE for the entire population. These figures show the individual MAPE for each area. To assist with visual comparison, only the MAPE for the 1st, 3rd and 5th forecasts are displayed, corresponding for August, October and December 2015. During the first month of the forecast, the autoregressive model produces an error of estimation below 10% for most of the areas. However, by the fifth forecast, the error almost doubles, resulting in almost 80% of the agents having more than 10% errors in estimations (see Fig. 16a, b and c). Thus, while the autoregressive model can capture the spatio-temporal nature of the PV adoption process, it is only able to produce short term forecasts.

The multivariable model reaches an error level of below 10% for most of the areas, during the first forecast period. Then, by the fifth forecast, the error significantly increases where almost 30% of the areas having more than 10% of errors. Nevertheless, the accumulation of errors over time reduces in comparison with the autoregressive model. These temporal results suggest that characterising the evolution of agents' heterogeneity improves the model's performance as they generate a lower MAPE. This suggests that the preferences of agents are dynamic and so is their decision-making. It could be argued that the PV autoregressive model disregards the evolution of the preferences and focuses only on the experiences and social dynamics, thus resulting in a larger accumulation of error.

5. Discussion

Despite ABMs' strength in providing insights on emergent system behaviour, two main limitations are still outstanding: the use of *rational choice*-based decision-making [33–35] and synthetic characterisation of their temporal dynamics, rather than utilising actual time horizons. Then, the use of ANNs as a basis for decision-making could be considered to help address the limitations of ABM. First, the model substitutes *rational choice*-based decision-making, adopting an experience-based approach. The model assumes that the agents' decision-making is driven by experience and perceptions [38–40] instead of complicated estimations about affordability [11], energy economics [16,62] or payback period [50,52,54], and rarely possessing perfect market information [17,33–35]. Secondly, the use of ANN provides an actual time horizon for the adoption process to take place. Moreover, given the need to reflect the evolution of the agents' preferences, the multivariable model updates the agents' characteristics at each time step during the simulation. As expected, the multivariable characterisation of the decision-making process improves the model's accuracy.

These autoregressive models can estimate the monthly PV adoption rate upmost of three months, with an accuracy higher than 80%. The results show that the autoregressive models perform better than the Bass model especially for the months in the middle of the time series, those with most of the fluctuations in the number of PVs. The Bass model minimises the overall estimation error, meaning that estimations with high accuracy will offset the estimations with low accuracy. Then, on average the model minimises the estimation error but has limited capabilities to adapt to temporal changes in the data behaviour.

The assumption within the autoregressive model is that the adoption process is driven by the experience and that exposure to PV technology is extended by including socioeconomic variables. Thus, multivariable models that integrates the income, household size and total electricity consumption variables improve the model performance from 90% to 95%. The multivariable model slows down the accumulation of the estimation errors, forecasting up to five months with an accuracy higher than 90%. This could be because the decision-making does not only consider the experience with the PV technology or the social influences but is also driven by the socioeconomic characteristics of the individuals [11]. Here the results show that a relatively small improvement in the model's performance leads to a lower accumulation of errors, suggesting that marginal changes in decision-making may result in more confident decision-making in the future.

Because of the data-driven nature of ANN included in the agent's decision-making, if the model is to include any data, for example the amount of apple consumption in a PC, it would be hard to explain the nature of the influence of such variables on the agents' preferences. In this context, this paper builds on the insights of the SR to make sense of the **electricity consumption** variables being selected, reflecting that those households with a high energy usage tend to be concerned about being self-sufficient [1] and reduce their energy bills [3]. This suggests that besides improving the performance of the model, the multivariable model can also identify some of the contexts of the adoption process. However, both ABM and ANN models developed in this paper may not be self-sufficient to provide a sense of meaning to the selected variables without the assistance of more contextual studies. For instance, the SR can translate the selected variables and the value of their coefficient to provide insights about the impact of each variable. Moreover, because the initial list of variables corresponds to a single relevant study [1], one can argue that a different list of variables may result in a different configuration and insights produced by the model.

The results presented in Section 4.1 suggest that the FiT incentive is more effective at the edges of the city than in the city centre. Therefore, it could be argued that a key difference is the density of residential buildings [1,55,57]. However, this variable was not selected during the step-wise process, whilst the multivariable model has already accounted for the effects of **income**, **household size** and **total electricity consumption**, which could be indirectly related to the density of residential buildings. Additionally, one can argue that the results and variables selected may be exclusive to the area of Birmingham and surrounding local authorities. Therefore, future analysis may include other cities of the UK to investigate whether the same list of variables produces similar results.

Drawing from the bounded rationality theory, an assumption within the model is that a group of individuals with similar socioeconomic characteristics make similar decisions. In which case, groups can be characterised as singular decision-making units. Then, this aggregated characterisation of agents allows the model to integrate the spatial dependence by weighting the spill-over effect against the distance between areas. Secondly, the peer-effect is captured by the influence that individuals within an area have over those in the same neighbourhood and thus reflects how individuals tend to associate with those of similar attitudes and values.

Although the spatial accuracy of the model inputs is important [61] gathering this information for entire populations is a challenge. Therefore, those ABMs that simulate data or use semi-empirical data, due to the high data-intensive needs, may present a loss in the spatial accuracy of results [14]. Alternatively, the ABM and ANN model defines agents as geographical areas, recognising that they have similar socioeconomic characteristics, common interests and presenting similar behaviours [38,45]. The use of spatially explicit data at the postcode district level allows the modelling of real-world layout, without excluding any locations. However, the aggregated characterisation is not without imperfections because of the aggregated nature of the data. There is potentially a loss of heterogeneity in the individuals' own decision choices in an area. The spatial resolution is driven by the data availability and temporal variability within it, highlighting a trade-off between these two elements. While higher spatial resolution improves spatial accuracy, this is at the expense of lower temporal variability in the adoption rates. Conversely, higher temporal resolution results in less data variability whilst the number of observations increase.

In a general sense, this paper develops an empirical characterisation of the Kunz [38] model and improves the Bierkandt et al. [45] model through the introduction of learning into the model. Furthermore, in the attempt to answer our research question, this work demonstrates that it is possible to explicitly characterise the spatio-temporal dynamics of the PV adoption. Such aggregated characterisation of the ABM allows for: (i) including the spill-over effect, which can be a proxy of the spatial dependence if it is weighted against the distance between areas; and (ii) implementing the peer-effect as the influence of past adoption rates within the same area.

5.1. Improving the agent's decision-making characterisation

The ABM and ANN model assumes that individuals with similar decision-making, attitudes and interests tend to spatially cluster [3], creating social networks [52] and that knowledge is generated through experience-based learning [38–40]. On the other hand, in conventional applications of ABMs, the agents are assigned with an adoption rule or adoption probability at the beginning of the simulation. Section 4.1 shows that early stages of the training produce large errors, as the model is not fed with enough data. Therefore, it follows that the actual behaviour in a specific location could still be unknown to the modeller (early stages of adoption). Therefore, it is argued that assigning an adoption rule to areas in the early stages of adoption may increase uncertainty.

Alternatively, the aggregated characterisation of the agents and the use of ANNs can allow the modelling of the entire population. Such an approach could disregard the specific individuals' preferences, by generating knowledge over time. The modelling approach takes into consideration historical data, even for those areas with a low number of PVs. This can be seen in the error produced by the model output, as the performance of ANN increases as it is fed with more data and starts adapting to the changes in data behaviours, i.e. as individuals have experienced more situations [65,67,68,83]. Additionally, the aggregated characterisation of the agents allows the ABM to provide a more realistic representation of the decision-making process. The characterisation of agents as geographical areas allows integrating the ANN as their decision-making, reducing the need for more detailed data at the individual agent level [93]. Besides, adjusting the strength of the spill-over effect as a function of the distance between the areas reflects the spatial dependence of the PV adoption. Thus, the model accounts for both the influence between and within locations [12]. The model aggregates the number of PVs registered by month, then in turn the social influence is defined by the number of PVs in the adjacent areas on monthly basis.

6. Conclusion

This paper explains the model design, then develops and empirically validates a novel spatio-temporal explicit agent-based model that integrates artificial neural networks into the agent's decision-making. The model advances ABMs by characterising the spatio-temporal dependence of the PV adoption process, whilst improving the agents' decision-making procedure. This approach not only builds on prior ABM research but draws from disciplines such as SR. The use of spatially explicit data sets allows reflection of individual behaviours for each location and follows the real-world layout for the city of Birmingham. The model utilises the ANN's capabilities to approximate historical PV data in the generation of knowledge and adapts to changes in data trends. Therefore, the model is shown to reduce uncertainties in the agents' decision-making.

The paper also shows how capturing the evolution of the population heterogeneity improves the model's predictive accuracy by introducing socioeconomic variables. In line with the literature, the results suggest that income is a key decision variable for households to adopt PV technology. Another important variable is electricity consumption indicating that households with high energy usage are more likely to adopt PVs due to potential concerns about being self-sufficient. The average household size variable captures the negative impact of PV ownership, as bigger families may have less cash flow for a PV installation. Although this multivariable characterisation does not increase the models' accuracy significantly (only 5%), it allows reducing drastically the errors from the forecast, from 10% to 2% for the first forecast and from 73% to 9% for the fifth forecast.

The results suggest that the model can account for the spatial, temporal and social dynamics that drive the adoption process. Furthermore, the ABM and ANN model can produce a better estimation than a Bass model, this could be because the Bass model does not consider the social effect nor the spatial dependence of the adoption process. In principle, spatio-temporally explicit forecasts could inform network planning and investment decisions of the energy industry. Yet, this potential currently is limited as the model is only able to produce short-term forecasts due to the limited availability of data for 60 time periods. Yet, with the increasing availability of data, this integrated approach has the potential to produce spatio-temporally explicit forecasts for longer periods.

CRediT authorship contribution statement

Ali Alderete Peralta: Conceptualization, Formal analysis, Methodology, Validation, Visualization, Writing – original draft. **Nazmiye Balta-Ozkan:** Conceptualization, Data curation, Resources, Supervision, Writing – review & editing. **Philip Longhurst:** Supervision, Writing – review & editing.

Data Access Statement

The data that support the findings of this study are openly available in Office for National Statistics Census Data at [https://www.ons.gov.uk/census/2011census] and Ofgem Feed-in Tariff Installation Report at [https://www.ofgem.gov.uk/publications/feed-tariff-installation-report-30-june-2021].

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Mexican National Council of Science and Technology scholarship (CONACYT). Partial research time for Balta-Ozkan is funded by UK Engineering and Physical Sciences Research Council award EP/R002320/1.

Appendix A. Summary of the potential socioeconomic variables

Figs. A1–A8.

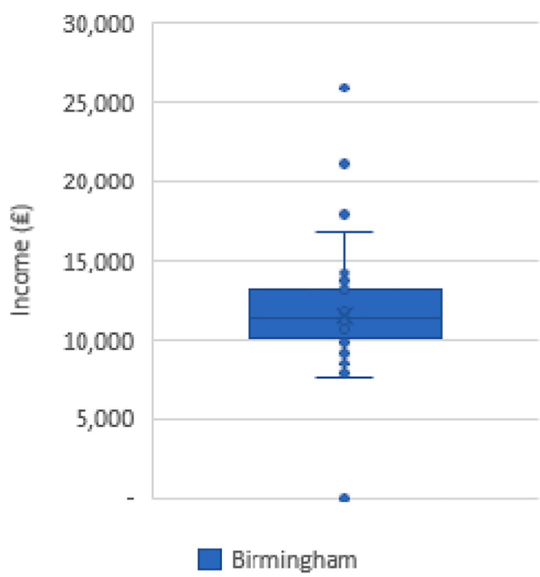

Fig. A1. Boxplot of the PC's income in Birmingham.

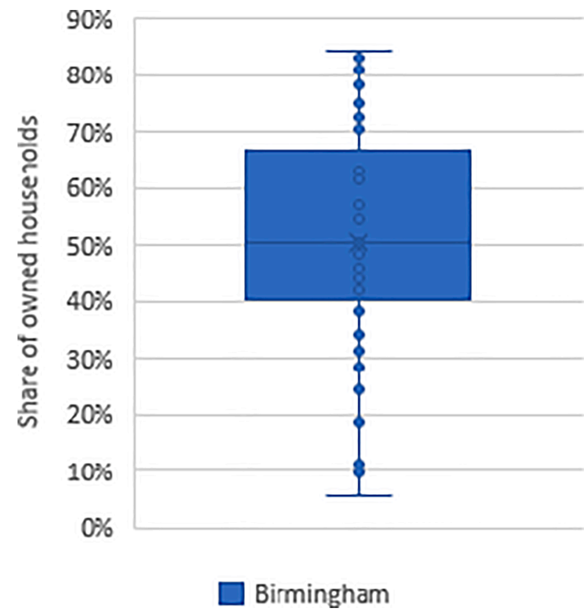

Fig. A3. Boxplot of the PC's share of owned households in Birmingham.

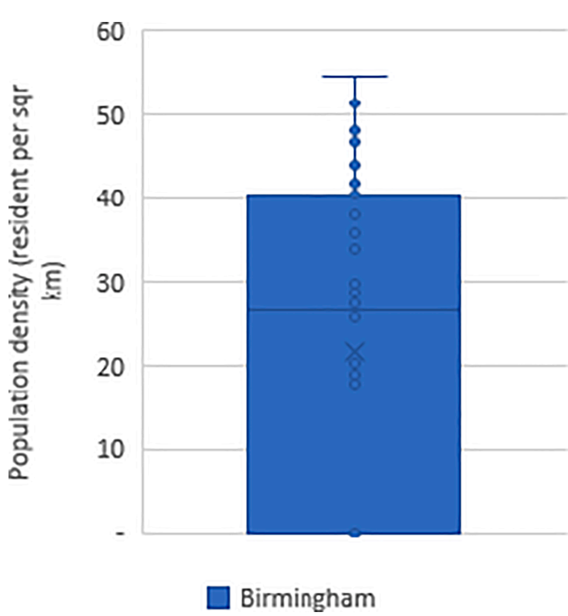

Fig. A2. Boxplot of the PC's population density in Birmingham.

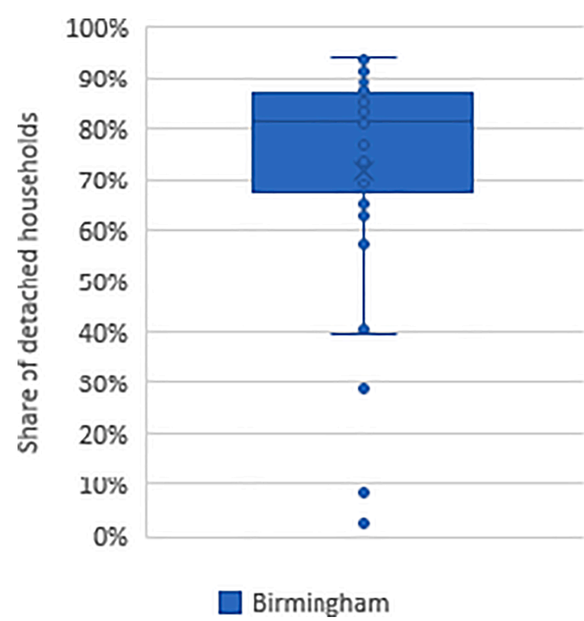

Fig. A4. Boxplot of the PC's share of detached households in Birmingham.

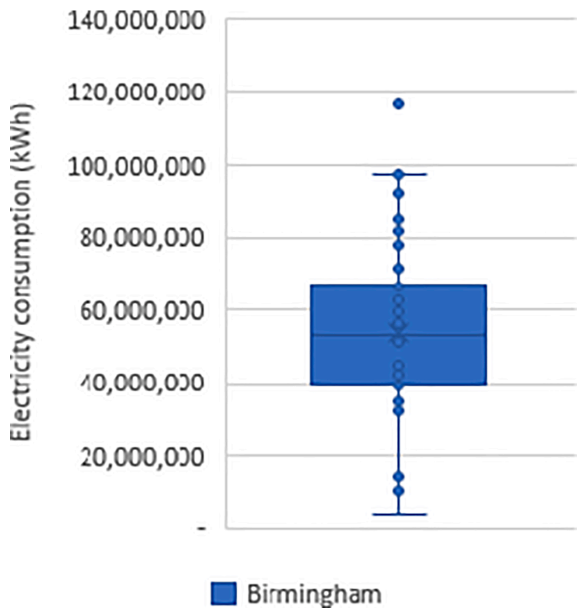

Fig. A5. Boxplot of the PC's electricity consumption in Birmingham.

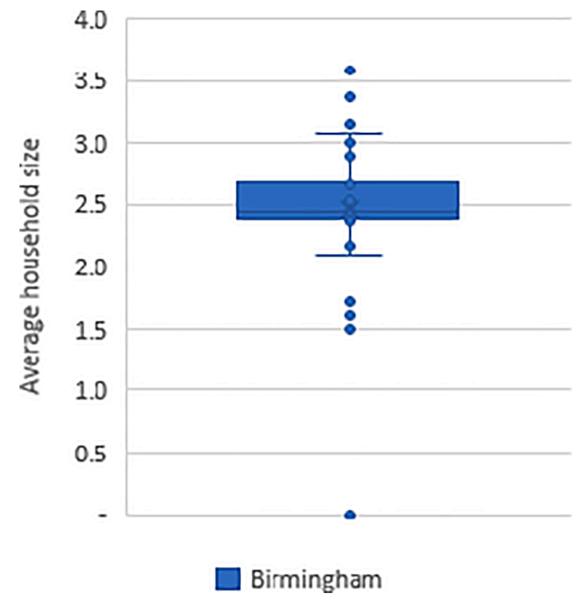

Fig. A7. Boxplot of the PC's average household size in Birmingham.

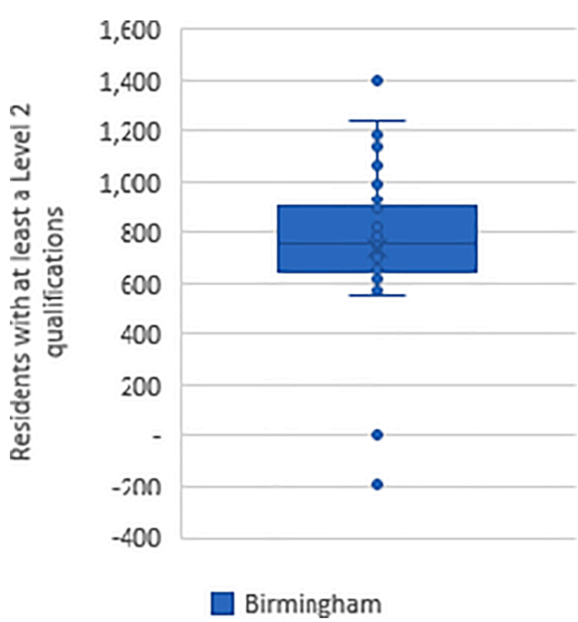

Fig. A6. Boxplot of the PC's education level in Birmingham.

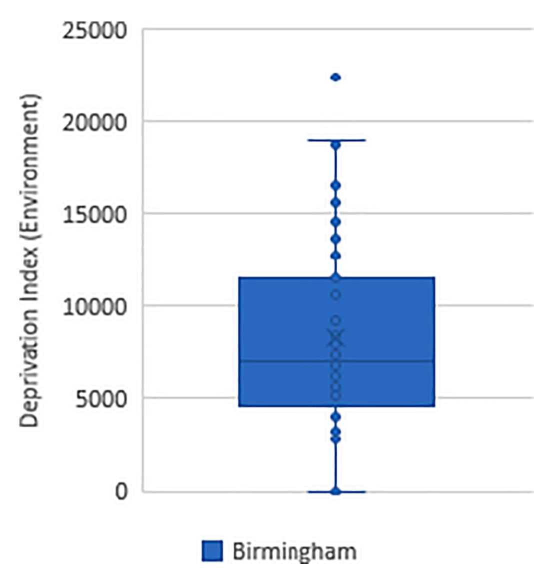

Fig. A8. Boxplot of the PC's Deprivation Index in Birmingham.

Appendix B. Bass model estimation

The Bass model is used to estimate the s-curve of each area, using Ordinary Least squares to estimates the parameter of the model. The model assumes 2050 as the horizon for total uptake of the PVs, and the number of residential building owned by the householder (per area) as the number of potential adopters. Then, following [46,79] the Bass model is defined as:

$$S(t) = \left(p + q^{*}\frac{Y(t)}{m}\right)(m-(Y(t).)$$

where

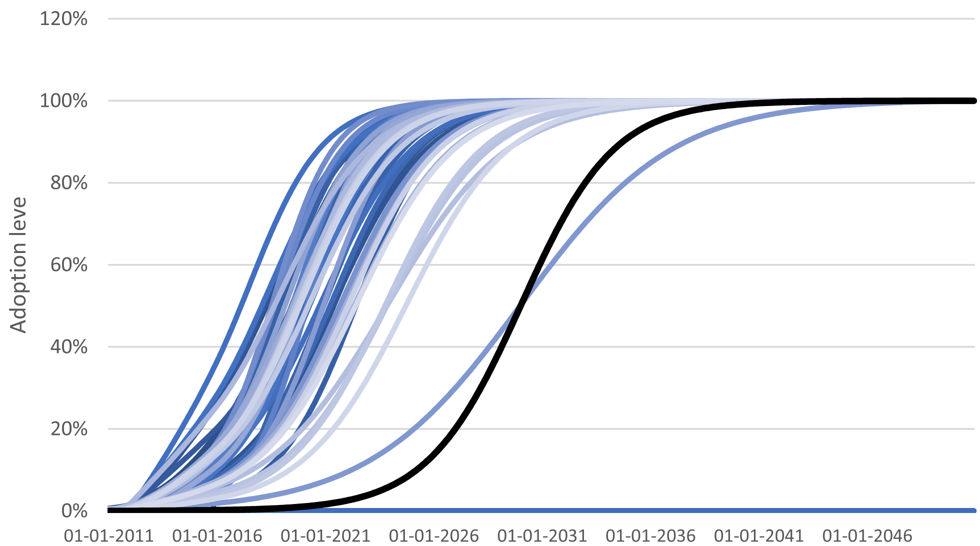

Fig. B1. Bass model's estimations for the PV uptake in the Birmingham's PCs.

S(t) is the number of new PV installations
p is the coefficient of innovation
q is the coefficient of imitation
Y(t) is the number of adopters
m is potential adopter

B 1shows the different S-curves for the PCs in Birmingham city, the line in black represent the estimation for the Total number of PVs in Birmingham (See Fig. B1).

Appendix C. Model implementation pseudocode

Algorithm 1 The PV adoption process

Agents characterisation

for each agent PC in Birmingham **do**
  PC.location ← actual population centroid
  $PC.PV_t$ ← PV installation dataset
  function AGENT_NEIGHBOURS()
    for each PC in agent_aux.Neighbours **do**
      agent_aux.calculatePVNeighbourhood()
      agent_aux.calculateDistance()
    end for
  end function
  function AGENT_ANN()
    for each PC in Birmingham **do**
      agent_aux.ANN(weight) ← random_between(0,1)
    end for
  end function
end for

Training

function TRAIN()
  for each agent PC in Birmingham **do**
    PC.train()
    PC.estimationError ← Mean absolute percentage error
  end for
end function

Forecasting

for each PC in Birmingham **do**
  $PC.PV_{t+1}$ ← PC.forecastPV()
  PC. calculatePVNeighbourhood()
end for

Appendix D. Estimation's convergence

See Fig. D1.

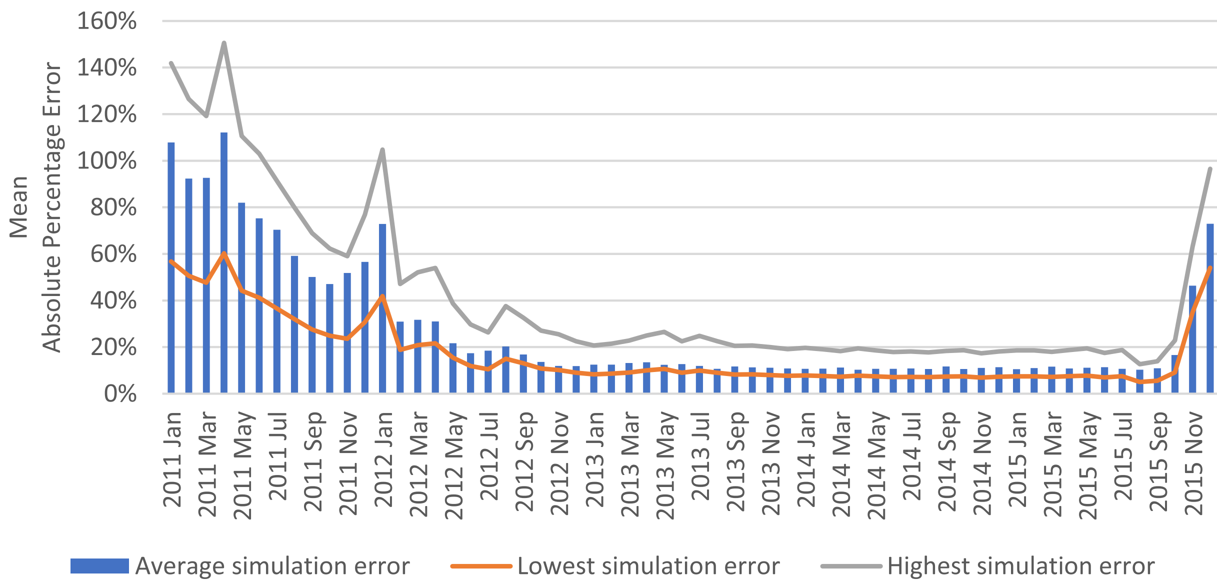

Fig. D1. Estimation error of the autoregressive model.

Appendix E. Training Algorithm

Algorithm 2 Backpropagation learning algorithm

Initialisation
1. **for each** agent PC in Birmingham **do**
2. PC.data ← historical PV installation data by month
3. **for each** month *m* in PC.data **do**
4. FORWARDS PASS
5. Starting from the input layer, use each activation function to compute the outputs
6. Use the synaptic weights to pass the outputs from each layer to the following one
7. Calculate the network output and the error of estimation
8. Check for stop condition
9. BACKWARDS PASS
10. Beginning from the last layer, compute the derivates of the output layer's function with respect to the estimation error
11. Compute the derivates of each other hidden layer with respect to the previous layer neurons function
12. Calculate the adjustment coefficient for each synaptic weight considering the previous layer neurons function
13. **end for**

end for

Appendix F. Stepwise process of variable selection

Table F1.

Table F1
Summary of the results from the stepwise process of variable selection.

MAPE	GDHI proxy	Population density	% Owned household share	% Detached household share	Electricity usage	Education level	Avg household size	Irradiation	CO2 proxy	Criteria
11.0%	Autoregressive model									
Step 1										
5.80%	X									Fixed
9.40%		X								Carry
8.60%			X							Carry
9.40%				X						Carry
9.00%					X					Carry
9.20%						X				Carry
8.80%							X			Carry
10.40%								X		Out
9.00%									X	Carry
Step 2										
6.2%	X	X								Carry
5.8%	X		X							Carry
6.2%	X			X						Carry
5.6%	X				X					Fixed
6.4%	X					X				Carry
5.6%	X						X			Fixed
6.6%	X								X	Out
Step 3										
5.8%	X	X			X					Carry
5.8%	X		X		X					Carry
6.6%	X			X	X					Out
5.8%	X				X	X				Carry
5.4%	X				X		X			Fixed
6.0%	X	X					X			Carry
6.4%	X		X				X			Carry
5.8%	X			X			X			Carry
6.8%	X					X	X			Out
Step 4										
6.0%	X	X			X		X			Out
6.2%	X		X		X		X			Out
6.2%	X				X	X	X			Out

References

[1] Balta-Ozkan N, Yildirim J, Connor PM. Regional distribution of photovoltaic deployment in the UK and its determinants: A spatial econometric approach. Energy Econ 2015;51:417–29. https://doi.org/10.1016/j.eneco.2015.08.003.
[2] Langheim R. Third Party-Owned PV Systems: Understanding Market Diffusion with Geospatial Tools. Energy Mark Predict Pap from 2014 AAAI Fall Symp Third. 2014.
[3] Graziano M, Gillingham K. Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environment. J Econ Geogr 2015;15:815–39. https://doi.org/10.1093/jeg/lbu036.
[4] Gupta R, Pena-Bello A, Streicher KN, Roduner C, Thöni D, Patel MK, et al. Spatial analysis of distribution grid capacity and costs to enable massive deployment of PV, electric mobility and electric heating. Appl Energy 2021;287:116504. https://doi.org/10.1016/j.apenergy.2021.116504.
[5] van der Kam MJ, Meelen AAH, van Sark WGJHM, Alkemade F. Diffusion of solar photovoltaic systems and electric vehicles among Dutch consumers: Implications for the energy transition. Energy Res. Soc Sci 2018;46:68–85. https://doi.org/10.1016/j.erss.2018.06.003.
[6] Barnacle M, Robertson E, Galloway S, Barton J, Ault G. Modelling generation and infrastructure requirements for transition pathways. Energy Policy 2013;52:60–75. https://doi.org/10.1016/j.enpol.2012.04.031.
[7] Müller J, Trutnevyte E. Spatial projections of solar PV installations at subnational level: Accuracy testing of regression models. Appl Energy 2020;265:114747. https://doi.org/10.1016/j.apenergy.2020.114747.
[8] Ipakchi A, Albuyeh F. Grid of the future. IEEE Power Energy Mag 2009;7:52–62. https://doi.org/10.1109/MPE.2008.931384.
[9] Ratnam EL, Weller SR. Receding horizon optimization-based approaches to managing supply voltages and power flows in a distribution grid with battery storage co-located with solar PV. Appl Energy 2018;210:1017–26. https://doi.org/10.1016/j.apenergy.2017.08.163.
[10] Morton C, Anable J, Yeboah G, Cottrill C. The spatial pattern of demand in the early market for electric vehicles: Evidence from the United Kingdom. J Transp Geogr 2018;72:119–30. https://doi.org/10.1016/j.jtrangeo.2018.08.020.
[11] Namdeo A, Tiwary A, Dziurla R. Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region. Technol Forecast Soc Change 2014;89:188–200. https://doi.org/10.1016/j.techfore.2013.08.032.
[12] Richter LL. Social Effects in the Diffusion of Solar Photovoltaic Technology in the UK. Cambridge Work Pap Econ 2013:40.
[13] Shafiei E, Thorkelsson H, Ásgeirsson EI, Davidsdottir B, Raberto M, Stefansson H. An agent-based modeling approach to predict the evolution of market share of electric vehicles: A case study from Iceland. Technol Forecast Soc Change 2012;79:1638–53. https://doi.org/10.1016/j.techfore.2012.05.011.
[14] Adepetu A, Keshav S, Arya V. An agent-based electric vehicle ecosystem model: San Francisco case study. Transp Policy 2016;46:109–22. https://doi.org/10.1016/j.tranpol.2015.11.012.
[15] Adjemian MK, Cynthia Lin CY, Williams J. Estimating spatial interdependence in automobile type choice with survey data. Transp Res Part A Policy Pract 2010;44:661–75. https://doi.org/10.1016/j.tra.2010.06.001.
[16] Sweda T, Klabjan D. An agent-based decision support system for electric vehicle charging infrastructure deployment. 2011 IEEE Veh. Power Propuls. Conf. VPPC 2011, IEEE; 2011, p. 1–5. https://doi.org/10.1109/VPPC.2011.6043201.
[17] Bale CSE, Varga L, Foxon TJ. Energy and complexity: New ways forward. Appl Energy 2015;138:150–9. https://doi.org/10.1016/j.apenergy.2014.10.057.
[18] Chaouachi A, Bompard E, Fulli G, Masera M, De Gennaro M, Paffumi E. Assessment framework for EV and PV synergies in emerging distribution systems. Renew Sustain Energy Rev 2016;55:719–28. https://doi.org/10.1016/j.rser.2015.09.093.
[19] Hughes N, Strachan N, Gross R. The structure of uncertainty in future low carbon pathways. Energy Policy 2013;52:45–54. https://doi.org/10.1016/j.enpol.2012.04.028.
[20] Ault G, Frame D, Hughes N, Strachan N. Electricity Network Scenarios for Great Britain in 2050; 2008.

[21] Kahraman C. Energy Management—Collective and Computational Intelligence with Theory and Applications. vol. 149. 2018. https://doi.org/10.1007/978-3-319-75690-5.
[22] Noonan DS, Hsieh L-HC, Matisoff D. Spatial Effects in Energy-Efficient Residential HVAC Technology Adoption. Environ Behav 2013;45:476–503. https://doi.org/10.1177/0013916511421664.
[23] Morton C, Wilson C, Anable J. The diffusion of domestic energy efficiency policies: A spatial perspective. Energy Policy 2018;114:77–88. https://doi.org/10.1016/j.enpol.2017.11.057.
[24] Moglia M, Cook S, McGregor J. A review of Agent-Based Modelling of technology diffusion with special reference to residential energy efficiency. Sustain Cities Soc 2017;31:173–82. https://doi.org/10.1016/j.scs.2017.03.006.
[25] Buchmann CM, Grossmann K, Schwarz N. How agent heterogeneity, model structure and input data determine the performance of an empirical ABM - A real-world case study on residential mobility. Environ Model Softw 2016;75:77–93. https://doi.org/10.1016/j.envsoft.2015.10.005.
[26] Macal CM, North MJ. Agent-based modeling and simulation. Proc. - Winter Simul. Conf., IEEE 2009:86–98. https://doi.org/10.1109/WSC.2009.5429318.
[27] Rai V, Robinson SA. Agent-based modeling of energy technology adoption: Empirical integration of social, behavioral, economic, and environmental factors. Environ Model Softw 2015;70:163–77. https://doi.org/10.1016/j.envsoft.2015.04.014.
[28] Robinson SA, Rai V. Determinants of spatio-temporal patterns of energy technology adoption: An agent-based modeling approach. Appl Energy 2015;151:273–84. https://doi.org/10.1016/j.apenergy.2015.04.071.
[29] Robinson SA, Stringer M, Rai V, Tondon A. GIS-Integrated Agent-Based Model of Residential Solar PV Diffusion. USAEE 2013, 2013.
[30] Macal CM, Graziano DJ, Ozik J. Modeling solar PV adoption: A social-behavioral agent-based framework. 2014 AAAI Fall Symp; 2014; FS: 24–31.
[31] Haifeng Z, Yevgeniy V, Letchford J, Lakkaraju K. Predicting Rooftop Solar Adoption Using Agent-Based Modeling. Energy Mark. Predict. Pap. from 2014 AAAI Fall Symp. Predict., 2014.
[32] Adepetu A, Keshav S. Understanding solar PV and battery adoption in Ontario. Proc. Seventh Int. Conf. Futur. Energy Syst. - e-Energy '16, New York, New York, USA: ACM Press; 2016, p. 1–12. https://doi.org/10.1145/2934328.2934333.
[33] Amblard F, Quattrociocchi W. Simulating social complexity: Social Networks and Spatial Distribution. In: Edmonds B, Meyer R, editors. Simulating Complex. second, 2018, p. 763–800.
[34] Farmer JD, Foley D. The economy needs agent-based modelling. Nature 2009;460:685–6. https://doi.org/10.1038/460685a.
[35] van Duinen R, Filatova T, Jager W, van der Veen A. Going beyond perfect rationality: drought risk, economic choices and the influence of social networks. Ann Reg Sci 2016;57:335–69. https://doi.org/10.1007/s00168-015-0699-4.
[36] Noori M, Tatari O. Development of an agent-based model for regional market penetration projections of electric vehicles in the United States. Energy 2016;96:215–30. https://doi.org/10.1016/j.energy.2015.12.018.
[37] Pahl-Wostl C. Agent based simulation in integrated assessment and resources management. Int. Environ. Model. Softw. Soc. (iEMSs 2002), vol. 229, 2002, p. 239–44.
[38] Kunz J. Group-Level Exploration and Exploitation: A Computer Simulation-Based Analysis. 2010:29:3 2011.
[39] Arthur WB. Inductive reasoning and bounded rationality (The El Farol problem). Am Econ Rev 1994;84:406–11. https://doi.org/10.2307/2117868.
[40] Simon HA. Bounded Rationality. Util. Probab., London: Palgrave Macmillan UK; 1990, p. 15–8. https://doi.org/10.1007/978-1-349-20568-4_5.
[41] Kang SG, Choi SH. Multi-Agent Based Beam Search for Real-Time Production Scheduling and Control. The Multi-Agent Based Beam Search Method, London: Springer London; 2013. https://doi.org/10.1007/978-1-4471-4576-9.
[42] Zhao T, Zhou Z, Zhang Y, Ling P, Tian Y. Spatio-Temporal Analysis and Forecasting of Distributed PV Systems Diffusion: A Case Study of Shanghai Using a Data-Driven Approach. IEEE Access 2017;5:5135–48. https://doi.org/10.1109/ACCESS.2017.2694009.
[43] Jupe SCE. Active Management of Distributed Generation based on Component Thermal Properties. Durham University, 2010.
[44] Masters CL. Voltage rise: the big issue when connecting embedded generation to long 11 kV overhead lines. Power Eng J 2002; 16: 5–12. https://doi.org/10.1049/pe:20020101.
[45] Bierkandt R, Wenz L, Willner SN, Levermann A. Acclimate—a model for economic damage propagation. Part 1: basic formulation of damage transfer within a global supply network and damage conserving dynamics. Environ Syst Decis 2014; 34: 507–24. https://doi.org/10.1007/s10669-014-9523-4.
[46] Al-Alawi BM, Bradley TH. Review of hybrid, plug-in hybrid, and electric vehicle market modeling Studies. Renew Sustain Energy Rev 2013;21:190–203. https://doi.org/10.1016/j.rser.2012.12.048.
[47] Samarasinghe S. Neural Networks for Applied Sciences and Engineering: From Fundamentals to Complex Pattern Recognition. Auerbach; 2016.
[48] Kök BV, Yilmaz M, Çakiroglu M, Kuloglu N, Sengür A. Neural network modeling of SBS modified bitumen produced with different methods. Fuel 2013;106:265–70. https://doi.org/10.1016/j.fuel.2012.12.073.
[49] Dematos G, Boyd MS, Kermanshahi B, Kohzadi N, Kaastra I. Feedforward versus recurrent neural networks for forecasting monthly japanese yen exchange rates. Financ Eng Japanese Mark 1996;3:59–75. https://doi.org/10.1007/BF00868008.
[50] Krebs F, Ernst A. A spatially explicit agent-based model of the diffusion of green electricity: Model setup and retrodictive validation. Adv. Intell. Syst. Comput., vol. 528, Springer, Cham; 2017, p. 217–30. https://doi.org/10.1007/978-3-319-47253-9_19.
[51] Schwarz N, Ernst A. Agent-based modeling of the diffusion of environmental innovations - An empirical approach. Technol Forecast Soc Change 2009;76:497–511. https://doi.org/10.1016/j.techfore.2008.03.024.
[52] Ernst A, Briegel R. A dynamic and spatially explicit psychological model of the diffusion of green electricity across Germany. J Environ Psychol 2017;52:183–93. https://doi.org/10.1016/j.jenvp.2016.12.003.
[53] Bale C, Gale W, McCullen N, Rucklidge A, Foxon T. Harnessing social networks for promoting adoption of energy technologies in the domestic sector. Energy Policy 2013.
[54] Krebs F. An empirically grounded model of green electricity adoption in Germany: Calibration, validation and insights into patterns of diffusion. JASSS 2017;20:10. https://doi.org/10.18564/jasss.3429.
[55] Davidson C, Drury E, Lopez A, Elmore R, Margolis R, W BR, et al. Modeling photovoltaic diffusion: an analysis of geospatial datasets. Environ Res Lett 2014;9:074009. https://doi.org/10.1088/1748-9326/9/7/074009.
[56] Dharshing S. Household dynamics of technology adoption: A spatial econometric analysis of residential solar photovoltaic (PV) systems in Germany. Energy Res Soc Sci 2017;23:113–24. https://doi.org/10.1016/j.erss.2016.10.012.
[57] Schaffer AJ, Brun S. Beyond the sun—Socioeconomic drivers of the adoption of small-scale photovoltaic installations in Germany. Energy Res Soc Sci 2015;10:220–7. https://doi.org/10.1016/j.erss.2015.06.010.
[58] Liu L, Zhao Y, Chang D, Xie J, Ma Z, Sun Q, et al. Prediction of short-term PV power output and uncertainty analysis. Appl Energy 2018;228:700–11. https://doi.org/10.1016/j.apenergy.2018.06.112.
[59] De Groote O, Pepermans G, Verboven F. Heterogeneity in the adoption of photovoltaic systems in Flanders. Energy Econ 2016;59:45–57. https://doi.org/10.1016/j.eneco.2016.07.008.
[60] Zhou Z, Zhao T, Zhang Y, Su Y. A data-driven approach to forecasting the distribution of distributed photovoltaic systems. Proc. IEEE Int. Conf. Ind. Technol., vol. 2018- Febru, IEEE; 2018, p. 867–72. https://doi.org/10.1109/ICIT.2018.8352292.
[61] Cui X, Kim HK, Liu C, Kao S-C, Bhaduri BL. Simulating the household plug-in hybrid electric vehicle distribution and its electric distribution network impacts. Transp Res Part D Transp Environ 2012;17:548–54. https://doi.org/10.1016/j.trd.2012.05.011.
[62] Eppstein MJ, Grover DK, Marshall JS, Rizzo DM. An agent-based model to study market penetration of plug-in hybrid electric vehicles. Energy Policy 2011;39:3789–802. https://doi.org/10.1016/j.enpol.2011.04.007.
[63] McCoy D, Lyons S. Consumer preferences and the influence of networks in electric vehicle diffusion: An agent-based microsimulation in Ireland. Energy Res Soc Sci 2014;3:89–101. https://doi.org/10.1016/j.erss.2014.07.008.
[64] Balta-Ozkan N, Watson T, Mocca E. Spatially uneven development and low carbon transitions: Insights from urban and regional planning. Energy Policy 2015;85:500–10. https://doi.org/10.1016/j.enpol.2015.05.013.
[65] da Silva IN, Hernane Spatti D, Andrade Flauzino R, Bartocci Liboni LH, dos Reis Alves SF. Multilayer Perceptron Networks. Artif. Neural Networks. Cham: Springer International Publishing; 2017. p. 55–115.
[66] Saraiva F de O, Bernardes WMS, Asada EN. A framework for classification of non-linear loads in smart grids using Artificial Neural Networks and Multi-Agent Systems. Neurocomputing 2015;170:328–38. https://doi.org/10.1016/j.neucom.2015.02.090.
[67] Lai LYWK. Basic Learning Principles of Artificial Neural Networks. Foreign-Exchange-Rate Forecast. With Artif. Neural Networks, Springer, Boston, MA; 2007, p. 27–37. https://doi.org/10.1007/978-0-387-71720-3_2.
[68] Hill T, Marquez L, O'Connor M, Remus W. Artificial neural network models for forecasting and decision making. Int J Forecast 1994;10:5–15. https://doi.org/10.1016/0169-2070(94)90045-0.
[69] Kaastra I, Boyd M. Designing a neural network for forecasting financial and economic time series. Neurocomputing 1996;10:215–36. https://doi.org/10.1016/0925-2312(95)00039-9.
[70] Werbos PJ. Backpropagation Through Time: What It Does and How to Do It. Proc IEEE 1990;78:1550–60. https://doi.org/10.1109/5.58337.
[71] Copiello S, Grillenzoni C. Solar Photovoltaic Energy and Its Spatial Dependence. Energy Procedia, vol. 141, Elsevier; 2017, p. 86–90. https://doi.org/10.1016/j.egypro.2017.11.017.
[72] Chamberlin G. Temporal Disaggregation; 2010.
[73] Liu L, Kong F, Liu X, Peng Y, Wang Q. A review on electric vehicles interacting with renewable energy in smart grid. Renew Sustain Energy Rev 2015;51:648–61. https://doi.org/10.1016/j.rser.2015.06.036.
[74] Raza MQ, Khosravi A. A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings. Renew Sustain Energy Rev 2015;50:1352–72. https://doi.org/10.1016/j.rser.2015.04.065.
[75] Lu H, Ma X, Ma M, Zhu S. Energy price prediction using data-driven models: A decade review. Comput Sci Rev 2021;39:100356. https://doi.org/10.1016/j.cosrev.2020.100356.
[76] Khosravi A, Nahavandi S, Creighton D. Quantifying uncertainties of neural network-based electricity price forecasts. Appl Energy 2013;112:120–9. https://doi.org/10.1016/j.apenergy.2013.05.075.
[77] Bodo P. MADness in the method: On the volatility and irregularity of technology diffusion. Technol Forecast Soc Change 2016;111:2–11. https://doi.org/10.1016/j.techfore.2016.05.011.
[78] Heymann F, Miranda V, Soares FJ, Duenas P, Perez Arriaga I, Prata R. Orchestrating incentive designs to reduce adverse system-level effects of large-scale EV/PV adoption – The case of Portugal. Appl Energy 2019;256:113931. https://doi.org/10.1016/j.apenergy.2019.113931.

[79] Linder S, Wirges J. Spatial diffusion of electric vehicles in the German metropolitan region of Stuttgart. ERSA Conf Pap 2011:1–27.
[80] Yamaguchi Y, Akai K, Shen J, Fujimura N, Shimoda Y, Saijo T. Prediction of photovoltaic and solar water heater diffusion and evaluation of promotion policies on the basis of consumers' choices. Appl Energy 2013;102:1148–59. https://doi.org/10.1016/j.apenergy.2012.06.037.
[81] Lee CY, Huh SY. Forecasting the diffusion of renewable electricity considering the impact of policy and oil prices: The case of South Korea. Appl Energy 2017;197: 29–39. https://doi.org/10.1016/j.apenergy.2017.03.124.
[82] Rogers EM. Diffusion of innovations. New York: Free Press; 2003.
[83] Zou J, Han Y, So SS. Overview of artificial neural networks. Artif. Neural Networks, vol. 458, Humana Press; 2008, p. 15–23. https://doi.org/10.1007/978-1-60327-101-1_2.
[84] Haykin S. Neural networks: a comprehensive foundation. 2nd ed.; 1999.
[85] Müller S, Rode J. The adoption of photovoltaic systems in Wiesbaden, Germany. Econ Innov New Technol 2013;22:519–35. https://doi.org/10.1080/10438599.2013.804333.
[86] DECC D of E and CC. Electricity System: Assessment of Future Challenges; 2012.
[87] Cherrington R, Goodship V, Longfield A, Kirwan K. The feed-in tariff in the UK: A case study focus on domestic photovoltaic systems. Renew Energy 2013;50:421–6. https://doi.org/10.1016/j.renene.2012.06.055.
[88] Burnett D, Barbour E, Harrison GP. The UK solar energy resource and the impact of climate change. Renew Energy 2014;71:333–43. https://doi.org/10.1016/j.renene.2014.05.034.
[89] Islam T, Meade N. The impact of attribute preferences on adoption timing: The case of photo-voltaic (PV) solar cells for household electricity generation. Energy Policy 2013;55:521–30. https://doi.org/10.1016/j.enpol.2012.12.041.
[90] Balta-Ozkan N, Yildirim J, Connor PM, Truckell I, Hart P. Energy transition at local level: Analyzing the role of peer effects and socio-economic factors on UK solar photovoltaic deployment. Energy Policy 2021;148:112004. https://doi.org/10.1016/j.enpol.2020.112004.
[91] Lin G, Zhang T. Loglinear residual tests of Moran's I autocorrelation and their applications to Kentucky breast cancer data. Geogr Anal 2007;39:293–310. https://doi.org/10.1111/j.1538-4632.2007.00705.x.
[92] Chen Y. Spatial autocorrelation approaches to testing residuals from least squares regression. PLoS ONE 2016;11:e0146865. https://doi.org/10.1371/journal.pone.0146865.
[93] Kyriakarakos G, Piromalis DD, Dounis AI, Arvanitis KG, Papadakis G. Intelligent demand side energy management system for autonomous polygeneration microgrids. Appl Energy 2013;103:39–51. https://doi.org/10.1016/j.apenergy.2012.10.011.