IMA Journal of Management Mathematics (2025) **00**, 1–29
https://doi.org/10.1093/imaman/dpaf019

Influence of consumer attitudes and social interactions in electric vehicle purchasing: integrating agent-based modelling and machine learning

Wen Xu[†] and Irina Harris
Cardiff Business School, Cardiff University, Aberconway Building, Colum Dr, Cardiff CF10 3EU, UK
[†]Corresponding author. Email: XuW20@cardiff.ac.uk

Jin Li
Cardiff School of Engineering, Cardiff University, Queen's Buildings, 14-17 The Parade, Cardiff CF24 3AA, UK

Peter Wells
Cardiff Business School, Cardiff University, Aberconway Building, Colum Dr, Cardiff CF10 3EU, UK

and

Gordon Foxall
Cardiff Business School, Cardiff University, Aberconway Building, Colum Dr, Cardiff CF10 3EU, UK,
Department of Business & Economics, Reykjavik University, Menntavegur 1, 102 Reykjavik, Iceland

[Received on 19 April 2024; accepted on 27 May 2025]

Accepted by: Prof. Aris Syntetos

Understanding consumer attitudes towards electric vehicle (EV) purchasing is essential for addressing the slow adoption rate. Traditional aggregated models of EV adoption employ a top-down approach, yet often fail to capture individual-level attitudes. In contrast, agent-based modelling (ABM) enables a bottom-up approach that reflects the heterogeneity in consumer decision-making and simulates social interactions. This study introduces an integrated model to analyze consumer attitudes towards EV adoption, incorporating empirical data and synthesized social interactions through ABM. The model undergoes micro-validation and optimization through parameter variation experiments and supervised machine learning (SML) methods. Results indicate that consumer attitudes towards EV purchasing are positively influenced by early adopters and environmental factors. These attitudes are further shaped by observing EVs in residential areas and receiving positive feedback from social circles. Perceptions of EVs as an environmentally friendly alternative also significantly enhance these attitudes. These findings suggest that marketers should develop targeted strategies for specific consumer segments, and policymakers should prioritize environmental awareness campaigns to drive positive public EV attitudes in the UK. This study emphasizes the importance of incorporating consumer heterogeneity and social interactions in attitude formation, which offers insights into EV promotion within Rogers's Diffusion of Innovations Theory.

Keywords: consumer attitudes; electric vehicle purchasing; social interaction; model micro-validation; model optimization; agent-based modelling.

1. Introduction

Electric vehicles (EVs), including battery, hybrid and plug-in hybrids, are integral to replacing traditional internal combustion engines (Onat *et al.*, 2014; van Staden *et al.*, 2024). The wide replacement aims to reduce greenhouse gas emissions, decrease fossil fuel reliance and advance sustainable transportation (He *et al.*, 2013; Sierzchula *et al.*, 2014; Xing *et al.*, 2024). Numerous countries have implemented initiatives to increase EV adoption through subsidies, expanded public charging infrastructure and R&D funding (Nie *et al.*, 2016; Song & Potoglou, 2020; Liu *et al.*, 2023). For example, the UK aims to ban new petrol and diesel car sales by 2030 and transition to zero-emissions vehicles by 2035 (DfT, 2020). However, EVs accounted for only around 11% of new car sales in 2020 (IEA, 2021). Similar trends in France and the Netherlands underscore the challenges in boosting EV adoption despite supportive policies (IEA, 2021).

Recent research highlights a gap between consumer attitudes, intentions and actual behaviour as a central factor limiting EV adoption (Lane & Potter, 2007; Martin *et al.*, 2016). This discrepancy arises partly due to perceived limitations in EV technology, including limited driving range (He *et al.*, 2014), high purchase costs (Kim *et al.*, 2020) and insufficient charging infrastructure (Berkeley *et al.*, 2018), all of which contribute to consumer hesitancy towards EV purchases. Conversely, studies show that consumer attitudes can be favourably influenced by factors like environmental awareness (Mairesse *et al.*, 2012) and hands-on experience, such as test drives (Schmalfuß *et al.*, 2017). These findings suggest that while technological and infrastructure-related barriers exist, targeted inventions focusing on education and experiential engagement can positively shift consumer perceptions of EVs.

Consumer attitudes play a crucial role in decision-making for innovation adoption (Rogers, 1976, 2003) and can reliably predict intentions and behaviours when measured accurately (Kokkinaki & Lunt, 1997). During the early stages of adoption, most consumers may hold negative attitudes towards EVs (Carley *et al.*, 2013; Kim *et al.*, 2016), making it challenging for policymakers and marketers to obtain accurate data on consumer intentions and behaviours. Understanding these attitudes is vital for designing strategies to foster early adoption and positively influence consumer perceptions.

Statistical analysis provides valuable insights into consumers' attitudes, intentions and actual purchasing behaviours. Traditional models often predict EV market trends from a top-down perspective (Gómez Vilchez *et al.*, 2013; Gómez Vilchez & Jochem, 2019), but they typically overlook the heterogeneity in individual decision-making (Agliari *et al.*, 2010; Shim *et al.*, 2018). This can result in overly optimistic forecasts (Wardle *et al.*, 2015; Wolinetz & Axsen, 2017). In contrast, agent-based modelling (ABM) captures consumer diversity (Eppstein *et al.*, 2011, 2015; Iftekhar *et al.*, 2011; Kiesling *et al.*, 2012) and simulates social interactions within artificial networks (Zhang & Vorobeychik, 2017), enabling detailed validation of individual attitudes and providing a comprehensive, bottom-up examination of complex consumer decision-making processes.

This article introduces a novel model to analyze consumer attitudes towards EV purchasing within a synthesized, socially interactive environment. By integrating ABM and SML, the model synthesizes new data from relevant theories and empirical sources to improve model performance. A key contribution of this study is the establishment of a framework for micro-validation and optimization of agent-based models through parameter variation experiments and SML methods. This model is applied to analyze consumer attitudes towards EV purchasing in Great Britain, utilizing empirical survey data (ONS, 2015).

The structure of this article is as follows. Section 2 introduces relevant theories and compares traditional modelling approaches with newer methods. Section 3 details the methodologies and data sources used in the study. Section 4 presents experiment results and discusses model performance. Section 5 concludes the study and outlines future research directions.

2. Theories and related literature

2.1. *Attitude-intention-behaviour theories*

Attitude, intention and behaviour are critical constructs in consumer decision-making. Fishbein & Ajzen (1975) posit that consumer attitudes towards behaviours precede both behavioural intentions and actual behaviour, measurable through salient beliefs. Fishbein (1979) later incorporated subjective norms in the Theory of Reasoned Action, while Ajzen (1985) extended this further, adding perceived behavioural control to develop the Theory of Planned Behaviour (TPB).

TPB views individual attitudes as largely static, overlooking the dynamic interplay between subjective norms and attitudes towards behaviour, a point critiqued by researchers, such as Prislin & Wood (2005), Lui *et al.* (2015) and Wan *et al.* (2017). Despite these critiques, studies show that consumer attitude significantly mediates the influence of subjective norms on intentions (Bananuka *et al.*, 2019), although they often fail to account for various social influences like descriptive norms (Cialdini *et al.*, 1990; Barth *et al.*, 2016). This limited view of complex social contexts contributes to the observed gap between attitudes and behaviours (Peattie, 2010; Axsen & Kurani, 2014) and has led to less emphasis on studying attitudes compared to intentions in predicting consumer purchasing behaviour.

Social influence theory (Kelman, 1958) acknowledges TPB by acknowledging that beliefs, attitudes, intentions and behaviours are shaped by social contexts. According to the theory, individuals tend to align their actions and beliefs with those of others (Prislin & Wood, 2005; Grindrod & Higham, 2012) and are influenced by various social norms, including descriptive, injunctive, subjective and provincial norms (Cialdini *et al.*, 1990; Castro-Santa *et al.*, 2023). For instance, consumer acceptance of EVs can increase if close friends hold favourable views EVs (subjective norm) or if EVs are commonly observed in their community (provincial norm), demonstrating the impact of peer perceptions and social contexts on individual decisions (Barth *et al.*, 2016).

Rogers's Theory of Diffusion of Innovations (Rogers, 1976, 2003) asserts that consumer attitudes are fundamental in shaping intentions and behaviours towards adopting innovations (Fry *et al.*, 2018). Rogers outlines a five-step decision-making process: (1) gaining initial awareness with little knowledge, (2) forming beliefs and attitudes through learning and persuasion, (3) developing an intention to adopt or reject, (4) trialling the innovation and (5) confirming the adoption decision. Variability in adoption outcomes is attributed to individual differences in initial conditions (e.g., innovativeness), characteristics (such as socioeconomic status) and perceptions of the innovation (e.g., its relative advantage) (Arts *et al.*, 2011).

2.2. *Aggregated and disaggregated models of innovation diffusion*

In innovation adoption research, the primary distinction between aggregated and disaggregated models, based on Bass's hazard model (Bass, 1969), lies in how they treat adoption behaviour across a population (Guseo & Mortarino, 2014). Aggregated models, such as system dynamics models, consider the entire population as a homogeneous group (Widiarta *et al.*, 2008). Adoption rates are treated as a function of time, generalizing the probability of adoption across the population (Mahajan & Muller, 1979). This approach enables a simplified, closed-form solution that captures the overall diffusion of innovation, accounting for external influences (e.g., advertising) and internal influences (e.g., word-of-mouth) under the assumption of uniform adoption tendencies (Goldenberg *et al.*, 2000; Kiesling *et al.*, 2012).

In contrast, disaggregated models, such as ABM, recognize heterogeneity by dividing the population into subgroups based on distinct demographic, psychographic or behavioural characteristics. This approach allows for variations in adoption likelihood among individuals, capturing differences due to factors such as age, socioeconomic status or lifestyle (An, 2012; Bell & Mgbemena, 2018; Baltas & Yannacopoulos, 2019; Foramitti *et al.*, 2024). While disaggregated models offer a more granular and accurate representation of diffusion patterns, they are computationally complex, as they often require modelling individual differences and interactions without a close-form solution (Garcia, 2005).

Building an agent-based model involves a complex calibration and validation process that lacks standardization (Gürcan *et al.*, 2013). The current practice typically uses black-box validation, in which models are evaluated by comparing the predicted adoption curves with actual adoption trends (Robinson, 2014). However, black-box validation faces two main criticisms. First, the limited availability of long-term historical data, which restricts the accuracy of adoption curves (Kiesling *et al.*, 2012; Zhang & Vorobeychik, 2017)—a particularly significant limitation in the early stages of innovation adoption when data is crucial for informing policy and marketing strategies. Second, there is a risk of overlooking Type I errors (false positives) when relying on predicted curves for model accuracy assessment (Nisbet *et al.*, 2009; Japkowicz, 2013).

Despite the advantages of ABM in capturing heterogeneous decision-making and simulating social interactions, its effectiveness depends on accurately replicating real-world decision-making processes (Jäger, 2019). Rand & Rust (2011) highlight the importance of micro-validation to ensure the reliability of agents' decision-making rules (DMR) within these models, emphasizing that detailed validation at the individual level is essential for producing robust, credible model outcomes.

2.3. *Inference of consumer attitude*

Consumer attitudes towards innovation can be inferred through their salient beliefs, as outlined by Fishbein & Ajzen (1975) and further supported by Barsyte & Fennis (2023). SML offers a data-driven approach for analyzing consumer attitudes, bypassing traditional requirements such as subject-matter assumptions in data collection, theoretical foundations for variable selection and detailed parameter interpretation (McFadden, 1981; Flynn *et al.*, 2014; Aboutaleb *et al.*, 2021; Moulaei *et al.*, 2024).

The development of an SML model involves three critical phases: training, validation and testing (Sarle, 2002). A high-fidelity SML model aims to minimize empirical risk, reduce overfitting and ensure generalizability to new datasets, with performance assessed through quantifiable metrics (Nisbet *et al.*, 2009). By effectively capturing and analyzing consumer beliefs and attitudes, SML provides a robust tool for inferring and predicting consumer responses to innovation.

2.4. *Integrating ABM and SML*

Integrating ABM and SML offers a robust approach to researching innovation adoption, leveraging their complementary strengths (Furtado, 2020). SML excels at identifying patterns through data-driven feature selection (Brathwaite *et al.*, 2017; Peysakhovich & Naecker, 2017; Abbasi *et al.*, 2024), while ABM captures complex interdependencies using precise rules that govern agent behaviours and social networks.

Recent efforts to integrate SML and ABM have led to three main approaches. First, SML classifiers inform agents' decision-making. For example, Zhang *et al.* (2016) used SML to model decision processes in a study on rooftop solar adoption, where agents' decisions were guided by data-driven classification. Similarly, Ravaioli *et al.* (2023) developed a data-driven agent-based model to explore the impact of policy incentives on agriculture land usage, with agents' decisions modelled solely by SML algorithms. Second, ABM-generated social interaction data enhances SML predictions. Hassouna (2012) exemplified this by using ABM to model social interactions and analyze their effects on customer retention in the UK mobile market. Zhou & Lund (2023) employed ABM to simulate interactions among prosumers, consumers, retailers and aggregators, examining the influence of stakeholder interactions on renewable energy adoption. Third, SML validates ABM outputs, as seen in Lamperti *et al.* (2017), who used machine learning for intelligent sampling in parameter selection to optimize ABM performance. Kotthoff & Hamacher (2022) applied a gradient-based SML model to validate ABM predictions on innovation diffusion, focusing on true positive rates to assess model accuracy.

Our research addresses three critical gaps in studying consumer attitudes towards innovation adoption.

(1) Undervaluation of consumer attitudes: despite the importance of consumer attitude in decision-making, it remains underutilized as an indicator of innovation adoption. This article introduces a novel framework that positions consumer attitudes as a central factor, linking attitudes to purchasing behaviour in a socially interactive environment.

(2) Need for micro-level validation in ABM: while ABM captures consumer heterogeneity, it often lacks micro-level validation, limiting its accuracy and applicability. Our approach addresses this gap by integrating ABM with SML, facilitating the micro-validation of agent DMR using empirical data.

(3) Incorporating social interactions and consumer beliefs: existing models frequently overlook the roles of social interactions and consumer beliefs in decision-making, both of which are essential for accurately simulating innovation adoption. This article contributes to this area by synthesizing social interactions and incorporating relevant consumer beliefs within our model, enhancing its predictive ability for consumer attitudes and adoption behaviour.

By applying this integrated ABM-SML framework to analyze consumer attitudes towards EV purchasing in Great Britain, using empirical survey data, we provide a robust, data-driven model that addresses these gaps and offers new insights into the role of attitudes in EV adoption.

3. Methodology

This study employs an SML approach to address two primary objectives: (1) to construct a Supervised Machine Learning—Attitude Classifier (SML-AC) for identifying respondent features associated with attitudes towards EV purchasing, and (2) to develop a Supervised Machine Learning—Parameter Validator (SML-PV) for validating and optimizing model parameters. An agent-based model was also developed, incorporating the SML-AC to define agents' DMR and the SML-PV to support model validation and optimization.

Section 3.1 introduces the primary data sources, Section 3.2 details the development of the SML-AC, Section 3.3 provides an overview of the ABM framework and Section 3.4 outlines the specifics of the SML-PV.

3.1. *Data*

The data for this study was sourced from the 'Public Attitudes Towards Electric Vehicles' survey, conducted as part of the Opinions and Lifestyle Survey by the UK Data Service (2014) and ONS (2015). This survey included 1,996 responses from individuals aged 16 and over living in Great Britain, with 962 samples collected in February 2014 and 1,034 in February 2015. Stratified random sampling based on household postcodes, coupled with face-to-face interviews, were the primary data collection methods.

After excluding unit non-responses and respondents unaware of EVs, the final valid sample sizes were 900 for 2014 and 943 for 2015.

To ensure compatibility with modelling requirements, data processing included adjustments to address potential biases due to questionnaire design features (e.g., skip logic, multiple-response questions, multipurpose design and non-substantive response options) (Vaus, 2002). Data preparation involved variables screening, recoding and dealing with missing values to avoid distorting the relationships between feature and target variables. The data preparation was verified by Pearson's test of independence.

The final dataset comprises 66 feature variables and one target variable. Feature variables are categorized as: (1) demographic and socio-economic characteristics—car ownership, driving licence ownership and travel frequency; (2) key considerations while buying a car—costs (purchase, fuel, maintenance, resale value, etc.), comfort, environment friendliness, electrically powered techniques, style/design, interior space, driving range, safety and speed; (3) concerns discouraging EV purchase—concerns like costs, limited model availability, lack of knowledge, battery range limitations, recharging constraints, safety, speed and perceived unreliability of electrically powered technology and (4) factors encouraging EV purchase—costs, model variety, environmental friendliness, effective battery range, recharging convenience, safety, vehicle size/aesthetics and perceived reliability of electrically powered technology. Most feature variables had unbalanced distributions in their specific categories, and the feature variables have unbalanced distributions across their categories, yet both datasets exhibited similar distributions of feature variables (DfT, 2016). For simplification, features listed in (1) are referred to as information features, while those in (2)–(4) are referred to as belief features.

The target variable has two categories: (1) a positive attitude towards buying EVs, including respondents who "already own EVs", "are thinking about buying EVs quite soon", "thinking about buying EVs but don't know when" and "thought about buying EVs but decided not to"; and (2) a negative attitude, including those who "have not really thought about buying EVs" and "do not drive". The target variable is unbalanced, with positive attitude corresponding to 19.6% of the 2014 data and 19.9% of the 2015 data. Both datasets were used to train, test, and validate the proposed SML classifiers and agent-based models.

The prepared survey datasets were imported into AnyLogic to construct consumer agents—900 agents for 2014 and 943 agents for 2015. Each consumer agent represents a survey respondent, carrying the same socio-demographic, attitudinal and behavioural characteristics as the corresponding respondent. To enhance model stability, a 4-fold replication of the respondent population was implemented, resulting in approximately 3,600 agents for 2014 and 3,772 agents for 2015 (see Section 3.3 for details).

3.2. *SML-AC*

To develop the SML-AC, we employed several machine learning techniques using SPSS Modeler 18.2.2 (IBM, 2018). Our primary objective was to ensure unbiased performance evaluation and optimal algorithm selection, achieved through K-fold cross-validation ($K = 5$). In this approach, the dataset D was divided into K equal-sized subsets $D_1, D_2, \ldots, D_K$, with $K - 1$ subsets used for training and the remaining subset used for testing, repeating this process across all folds (Clark, 2003; Ozdemir, 2016). The cross-validation error was calculated as follows:

$$\text{Cross Validation}_{\text{error}} = \frac{1}{K} \sum_{i=1}^{K} \text{error}\left(D_i^{\text{test}}\right) \tag{1}$$

where error (D_i^{test}) represents the classification error for the ith fold. Reducing this error minimizes the risks of overfitting and selection bias.

To identify relevant features for predicting attitudes toward EV adoption, we used a filter-based feature selection method, applying Pearson's chi-square test of independence (Fishbein & Ajzen, 1975; Ranganathan *et al.*, 2017). For each feature i, the chi-square statistic χ_i^2 was calculated as:

$$\chi_i^2 = \sum \frac{(O_i - E_i)^2}{E_i} \tag{2}$$

where O_i is the observed frequency and E_i is the expected frequency. This allowed us to rank features by their statistical association with the target variable and eliminate features with low correlation. In our experiment, features with a Pearson's chi-square value higher than 0.95 were selected for subsequent modelling, indicating a significant correlation between the feature and the target.

To address the issue of data imbalance, we applied the Synthetic Minority Oversampling Technique (SMOTE) with the M-nearest neighbours algorithm. SMOTE generates synthetic samples of the minority class by interpolating between existing instances and their nearest neighbours, as follows:

$$X_{\text{new}} = X_i + \lambda \left(X_j - X_i \right) \tag{3}$$

where X_i is a minority class sample, X_j is one of its nearest neighbours and λ is a random number between 0 and 1. This approach improves training data balance without introducing noise (Chawla *et al.*, 2002; Luengo *et al.*, 2010). Our experiment set the value of M to 5, a commonly used value recommended by Pertiwi *et al.* (2020). By applying SMOTE to our data, 176 samples with a positive EV attitude were synthesized.

We applied several classification algorithms, including logistic regression, Classification and Regression Tree (CART), Chi-Square Automatic Interaction Detector (CHAID) and C5.0 decision tree. These algorithms were chosen due to their effectiveness with categorical variables, aligning with the structure of the ONS datasets (Song & Lu, 2015; Hoffmann, 2016). More importantly, these algorithms offer a clear explanation of the relationship between independent and dependent variables, which is essential for providing insights into how independent variables influence the dependent variable in ABM simulations. While more complex algorithms, such as support vector machines and neural networks, could have been used, they were not selected because they do not provide a transparent understanding of these relationships (Clark, 2003).

The logistic regression model is defined as

$$P\,(\text{Positive EV attitude}) = {}^{1}\!\big/_{1 + e^{-(\beta_0 + \sum_{i=1}^{n} \beta_i X_i)}} \tag{4}$$

where $P\,$(Positive EV attitude) is the probability of an agent holding a positive attitude towards EV purchasing, X_i are predictor variables (e.g., attitudinal beliefs and demographic factors) and β_i are coefficients estimated from the data.

The decision tree algorithms (CART and C5.0) split the dataset into homogenous subgroups based on predictor variable values. For CART, the splitting criterion is based on Gini impurity

$$\text{Gini}(D) = 1 - \sum_{i=1}^{C} p_i^2 \tag{5}$$

where p_i is the proportion of instances of class i in dataset D and C is the number of classes. Each node splits to maximize the reduction in Gini impurity, refining the homogeneity of the subgroups.

The models were evaluated using K-fold cross validation (as described above) to calculate performance metrics, such as sensitivity, specificity and G-mean (Nisbet *et al.*, 2009). Sensitivity and specificity are defined as follows:

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \tag{6}$$

$$\text{Specificity} = \text{TN}/(\text{TN} + \text{FP}) \tag{7}$$

where TP is true positive, TN is true negative, FP is false positive and FN is false negative. The G-mean, representing the geometric mean of sensitivity and specificity, is calculated as

$$G - \text{mean} = \sqrt{\text{Sensitivity} * \text{Specificity}}. \tag{8}$$

This ensures that the classifier is assessed not only on its accuracy in identifying positive cases but also on its ability to minimize false positives. We evaluated and compared three decision tree algorithms and logistic regression on the 2014 data, followed by further testing on the 2015 data to ensure generalizability.

3.3. *Agent-based modelling*

An ABM was developed to (1) synthesize social interaction data as agent-level social interaction features and (2) simulate consumer attitudes towards EV adoption based on synthesized social activities and DMR.

3.3.1. *Agent initialization.* This agent population represents a microcosm of the ONS dataset, with each characteristic's distribution designed to match the empirical data. Probability distribution functions for each characteristic were defined based on the ONS data, ensuring that both population-wide and individual-level characteristics were realistic. Mathematically, each agent A_i is initialized as a vector using the following equation:

$$A_i = \left[\text{age}_i, \text{gender}_i, \text{education}_i, \text{car}_{\text{ownership}_i}, \text{belief}s_i\right]. \tag{9}$$

3.3.2. *Synthesis of agent-agent interaction.* Agent-agent social interactions were synthesized by assigning each agent to a social network (Hamill & Gilbert, 2009, 2016). Each agent was placed within a social circle composed of close friends, modelled using a distance-based network in AnyLogic (version 8.7.3 University). Let $G = (V, E)$ represent the agent network, where V is the set of agents and E is the set of edges denoting bidirectional links between agents. A link was established between two agents A_i and A_j if the Euclidean distance $d(A_i, A_j)$ was below a predefined threshold $d_{\max}$, mathematically defined as

$$d\left(A_i, A_j\right) \leq d_{\max}. \tag{10}$$

These social links allowed agents to communicate their attitudes with close friends, facilitating Discussions with Friends (DF) interactions. Attitude changes resulting from DF interactions were modelled using a social influence model (Barth *et al.*, 2016; Hamill & Gilbert, 2016; Bennett & Vijaygopal, 2018). When an agent A_i's attitude was influenced by the attitudes of their social circle, the change in attitude was governed by the following function:

$$\widehat{\text{Attitude}}_{\text{DF}} = f\left(\alpha_{\text{DF}}, \widehat{\text{Attitude}}_{\text{prev}}, \text{friends\_attitudes}\right) \tag{11}$$

where α_{DF} represents the influence coefficient, capturing the strength of social influence from friends. $\widehat{\text{Attitude}}_{\text{prev}}$ is the agent's prior attitude. friends_attributes is the aggregated attitude of the agent's close friends. This setup allowed for realistic simulation of how peer influences shape individual attitudes, particularly in the context of consumer attitudes towards EV adoption.

The potential for attitude change was modelled probabilistically using a uniform random variable. The probability of an agent being influenced by a DF interaction, denoted P_{DF}, was calculated as follows:

$$P_{\text{DF}} = \text{uniform}\,(0, 1) < \text{DFprobability} \tag{12}$$

where DFprobability controls the likelihood of an attitude change due to DF interaction. This parameter was initially assigned a value referring to the Amazon Mechanical Turk (AMT) survey (Krupa *et al.*, 2014) (see Table 1 below) to enable model development and testing. The value was then calibrated through sensitivity analysis (i.e., parameter variation experiment in AnyLogic) to align with the attitudinal patterns observed in the ONS survey. Additionally, demographic distributions of respondents from the ONS and AMT surveys were compared in Appendix C, indicating similarities in gender and income distributions.

Referring to AMT survey findings, we assumed that respondents in the AMT survey shared similar sensitivities to social interactions as subgroups with comparable attitudes towards EV purchasing in the ONS data. This assumption provided an initial parameter value for configuring and testing the model, with values subsequently validated and optimized using empirical data in the final experiments. This method ensures that the assumption does not introduce bias into the experimental results.

3.3.3. *Synthesis of agent-environment interaction.* In addition to social interactions, agents could be influenced by their environment through the observation of EVs on roads (OR interaction). Each agent was assigned to an artificial residential environment based on geographical information from the ONS. Within this environment, agents could observe EVs in their vicinity, with the OR interaction as a provincial norm effect (Barth *et al.*, 2016).

The attitude changes due to OR interaction was modelled using a social infection model (Hamill & Gilbert, 2016). For each agent, the likelihood of being influenced by observing EVs was calculated as

$$P_{\text{OR}} = \text{uniform}\,(0, 1) < \text{ORprobability} \tag{13}$$

TABLE 1. DFprobability *and* ORprobability*, based on* Krupa et al. (2014)

AMT data			Agents based on ONS data		
Attitude of respon-dents	**Probability of attitude change due to DF**	**Probability of attitude change due to OR**	**Attitude of agents**	**Initial value for DF probability**	**Initial value for OR probability**
Would consider buying EVs	28.3%	12.7%	Already own EVs Thinking about buying EVs quite soon	28.3%	12.7%
Might consider buying EVs	24.7%	9.3%	Thinking about buying EVs but do not know when Thought about buying EVs but decided not to	24.7%	9.3%
Would not consider buying EVs	20.4%	5.7%	Have not thought about buying EVs Do not drive/do not need a car	20.4%	5.7%

where ORprobability represents the probability that an agent's attitude shifts due to the observation of EVs on roads. This parameter was fine-tuned through sensitivity analysis, drawing on observed patterns from the AMT survey. Table 1 below provides an overview of DFprobability and ORprobability values based on Krupa *et al.* (2014), linking the AMT-reported probabilities of attitude changes due to DF and OR interactions with corresponding agent attitudes.

If the condition is met, the agent's new attitude, denoted as $\widehat{\text{attitude}}_{\text{OR}}$, is influenced by the proportion of EVs within the residential environment. Mathematically, the change in attitude is represented as

$$\widehat{\text{Attitude}}_{\text{OR}} = f\left(\alpha_{\text{OR}}, \widehat{\text{Attitude}}_{\text{prev}}, \text{observed\_EVs}\right) \tag{14}$$

where α_{OR} is the OR interaction coefficient, reflecting the strength of environmental influence. observed_EVs represents the number of EVs observed within the agent's environment.

This configuration allows agents to adjust their attitudes in response to the visibility of EVs in their local surroundings, capturing the influence of environmental exposure on consumer attitudes towards EV adoption. By calibrating the OR interaction parameter, the model can accurately reflect the influence of observed EV density, adding realism to the simulation of provincial norms and environmental effects.

3.3.4. *Agent DMR with social interactions.* In our agent-based model, agent activities were governed by DMR that integrated both social interactions (DF and OR) and outputs from the SML-AC classifier. These rules determined whether an agent's attitude would shift based on social interactions.

For instance, with the DMR incorporating DF interaction, denoted as $\text{DMR}_{\text{LR+DF}}$, expressed the agent's attitude as

$$\text{Attitude}\widehat{_{\text{DMR\_LR+DF}}} = f\left(\alpha_{\text{DF}}, \widehat{Attitude}_{DMR\_LR}\right) \tag{15}$$

where $\widehat{\text{Attitude}}_{\text{DMR\_LR}}$ is the agent's attitude predicted by the logistic regression classifier (that is SML-AC), and α_{DF} is the coefficient representing the influence of DF interaction.

Similarly, the DMR incorporating OR interaction, denoted as $\text{DMR}_{\text{LR+OR}}$, was governed by

$$\widehat{\text{Attitude}}_{\text{DMR\_LR+OR}} = g\left(\alpha_{\text{OR}}, \widehat{\text{Attitude}}_{\text{DMR+LR}}\right). \tag{16}$$

These DMRs were evaluated by comparing the predicted agent attitudes against the actual observed attitudes from the ONS data. To optimize model performance, sensitivity analysis was conducted by varying the parameters α_{DF} and α_{OR}, allowing for fine-tuning of social influence effects on agent decision-making.

3.4. *SML-PV*

The parameters DFprobability and ORprobability represent the likelihood of agents changing their attitudes due to the DF and OR social interactions, respectively. These parameters capture the heterogeneity in agents' social interactions and play a crucial role in shaping the simulated patterns of agent attitudes across the population. To determine the optimal values, we conducted parameter variation experiments by systematically varying these parameters and assessing to evaluate the model's performance.

The experiments involved varying DFprobability and ORprobability from 0 up to twice the values in Krupa *et al.* (2014), with increments set as one-quarter of the final value. For each set of parameter values, the model's performance was evaluated using G-mean (see (8) in Section 3.2). The goal was to optimize these metrics by identifying parameter values that maximized both the model's ability to correctly classify both positive and negative cases.

To further analyze, the results of the parameter variation experiments and identify the optimal parameters, an SML-based parameter validator (SML-PV) was constructed. The following techniques were applied to build the validator: 5-fold cross-validation (see (1) in Section 3.2) for unbiased performance evaluation, SMOTE with M-nearest neighbours algorithm (see (3) in Section 3.2) to address data imbalance, logistic regression (see (4) in Section 3.2) and decision trees (C5.0 and CART, see (5) in Section 3.2) for algorithm comparison.

The SML-PV enabled a structural exploration of parameter spaces, enhancing model robustness by ensuring the selected parameters align with empirical data while achieving optimal classification performance.

3.5. *SML-ABM approach for the integrated model*

The model integrating SML and ABM consists of five functional modules (Fig. 1).

(1) *SML-AC (attitude classifier):* the first module involved constructing a logistic regression classifier (as described in Section 3.2) to identify key features associated with respondents' attitudes towards EV adoption.

(2) *Synthesis of social interaction features*: the second module synthesized social interaction features, which were not present in the ONS datasets, by using ABM and incorporating insights from the AMT datasets as along with Social Influence Theory. These synthesized features enabled realistic social dynamics within the agent population.

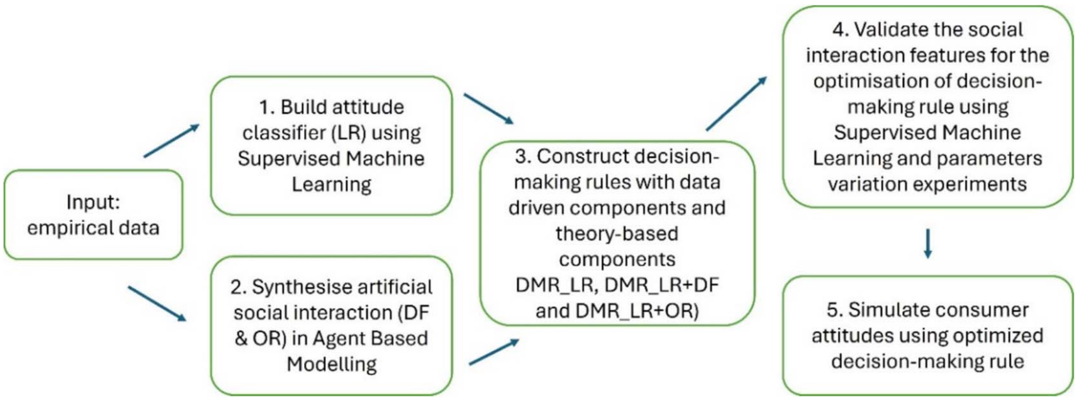


FIG. 1. Model structure.

(3) *DMRs for attitude formation*: in the third module, the synthesized social interaction features were combined with the logistic regression output from the SML-AC to formulate agents' DMRs for attitude formation.

(4) *Parameter variation with SML-PV*: This module involved assessing these DMRs through parameter variation experiments in the ABM. Optimized parameters for social interactions were selected using a C5.0 decision tree classifier, as outlined in the SML-PV (Section 3.4).

(5) *Simulation of consumer attitudes*: Finally, the agent-based model, now equipped with the optimized social interaction parameters, was used to simulate consumer attitudes towards EV adoption.

The integrated SML-ABM approach allowed for data-driven exploration of consumer attitudes, providing a robust framework to simulate and predict the impact of social and environmental influences on EV adoption.

4. Integrated model and results

4.1. *SML-AC classifier for consumer attitude*

To build the SML-AC, logistic regression, CHAID, CART, and C5.0 decision tree algorithms were initially selected. The performance of these algorithms was evaluated using several metrics, shown in Table 2, based on 5-fold cross-validation experiments. Logistic regression was selected as the final model due to its superior performance in key metrics, such as the area under the receiver operating characteristic curve (AUC), sensitivity and *G*-mean, with a low standard deviation (SD). Compared to other classifiers, logistic regression demonstrated the highest mean sensitivity, which is crucial for accurately identifying consumers with a positive EV attitude, ensuring the accuracy of subsequent ABM simulations. While CHAID, CART and C5.0 were also tested, they showed either lower means or higher SDs in sensitivity and specificity, indicating less stability in classification. Specifically, CHAID exhibited a high SD in specificity (7.4%), suggesting instability in classifying negative attitudes. CART and C5.0 also displayed high SDs, further supporting the choice of logistic regression. When trained on the ONS 2014 data and tested on the 2015 data, the logistic regression classifier achieved a sensitivity of 70.74%, specificity

TABLE 2. *Metrics for SML-AC, 5-fold cross validation experiment*

SML-AC candidates	**AUC (Mean, SD)**	**Sensitivity (Mean, SD)**	**Specificity (Mean, SD)**	***G*-mean (Mean, SD)**
CHAID	(70.53%, 2.29%)	(70.72%, 3.51%)	(62.86%, 7.40%)	(66.52%, 2.99%)
Logistic regression	(71.30%, 3.52%)	(70.11%, 3.77%)	(59.42%, 3.54%)	(64.48%, 1.56%)
CART	(66.10%, 3.83%)	(64.24%, 7.16%)	(62.82%, 10.78%)	(63.11%, 4.23%)
C5.0	(66.85%, 3.59%)	(53.42%, 14.0%)	(70.83%, 8.53%)	(60.63%, 6.42%)

of 61.46% and a *G*-mean of 65.94%, confirming its generalizability. Additionally, logistic regression's computational efficiency and ability to handle categorical data without overfitting made it the preferred model for this work.

Table 3 presents the B parameter (B), standard error (S.E.), Wald statistic, degrees of freedom (df), significance level (Sig.) and odds ratio (Exp(B)) for the predictors in the logistic regression classifier. Results indicate that three belief factors and two demographic features significantly affect consumer attitudes (Sig < 0.05). Belief factors included the importance of environmental friendliness, the significance of electrically powered technologies and concerns about limited battery range. For example, respondents who rated 'environmental friendliness as important when buying a car or van' were more likely to have a positive attitude towards EV purchasing compared to those who responded 'do not know' (the reference category). Demographically, car ownership and the highest level of qualification also played significant roles.

The logistic regression classifier labelled 42% of the 2014 respondents as having positive attitudes—higher than the actual figure of 19.6%—leading to over-optimistic predictions. This difference stemmed from sampling biases in the imbalanced data on consumer attitudes during early stages of innovation adoption (Rogers, 2003; Kiesling *et al.*, 2012), resulting in a Type I error, where many negative cases were misclassified as positive.

While undersampling is commonly used to improve classifier performance on unbalanced datasets, it may lead to data loss and potential misinterpretations (Krupa *et al.*, 2014; Mao *et al.*, 2021). An alternative approach involves adding effective predictors to the classifier, which can improve accuracy. Here, ABM was used to synthesize social interaction features, which may lead to more accurate classification of consumer attitudes (Barth *et al.*, 2016; Gallagher *et al.*, 2018).

4.2. *Synthesis of social interaction features using ABM*

An agent-based model was constructed to synthesize social interaction features based on the ONS data, by situating agents within a distance-based social network and a modelled residential environment. Agents engaged in predefined social activities, allowing them to interact with their social circles (DF interaction) and environment (OR interaction) through the observation of EVs on roads. These interactions were integrated into a social interaction component, augmenting the agents' original DMRs (as described in Section 3.3). The performances of the augmented rules, DMR_LR+DF and DMR_LR+OR, were then evaluated.

The synthesized social interaction features displayed similar patterns to the societal attitudes of AMT respondents towards EV adoptions (Krupa *et al.*, 2014). Figure 2 presents the population distributions of AMT respondents and simulated agents whose attitudes towards EV adoption were influenced by social

TABLE 3. *Predictors and parameters of logistic regression classifier*

Predictors and parameters in the selected SML-AC

	B	S.E.	Wald	df	Sig.	Exp(B)
Environmental friendliness as an important factor			22.201	2	0.000	
No	3.169	1.075	8.698	1	0.003	23.784
Yes	3.654	1.072	11.621	1	0.001	38.648
'Do not know or refusal (spontaneous only)' is coded as the reference category for 'environmental friendliness as an important factor'						
Electrically powered technologies as an important factor			19.342	1	0.000	
No	−1.272	.289	19.342	1	0.000	0.280
'Do not know or refusal (spontaneous only)' and 'yes' are combined due to their small proportions (8.4% and 5.4%, respectively), and coded as the reference category for 'electrically powered technologies as an important factor'						
Limited battery range as a deterrent			16.537	2	0.000	
No	.986	.326	9.168	1	0.002	2.680
Yes	1.309	.335	15.245	1	0.000	3.701
'Do not know or refusal (spontaneous only)' is coded as the reference category for 'limited battery range as a deterrent'						
Car ownership			40.259	3	0.000	
None	−.472	.348	1.842	1	0.175	0.624
One car	.600	.300	4.017	1	0.045	1.823
Two cars	.953	.311	9.418	1	0.002	2.594
'Three or more cars' is coded as the reference category for 'car ownership'						
Highest level of qualification			64.652	3	0.000	
Degree of equivalent	1.978	.280	49.829	1	0.000	7.230
Below degree level	1.062	.273	15.097	1	0.000	2.891
Other qualifications	1.557	.318	23.942	1	0.000	4.746
'None (no formal qualifications)' is coded as the reference category for 'highest level of qualification'						
Constant	−5.009	1.109	20.413	1	0.000	0.007
Sig. was calculated at the significant level of 0.05						

interaction activities, namely 'discussions with close friends' and 'observation of EVs on roads'. These results indicate that the simulated social interactions are compatible with empirical data, validating their application in our model to investigate the impact of social interactions on consumers' attitudes towards EV adoption.

Figure 3(A–C) illustrates the model performance for individual agent attitude classification under three rules: DMR_LR, DMR_LR+OR and DMR_LR+DF. The evaluation metrics were averaged over 30 replications to achieve statistically stable results. Paired sample *t*-tests across different rules indicated statistically significant differences among all test groups. Figure 3(D) and (E) shows that a high rank indicates that the predicted attitude ratio is closer to empirical data, while a low rank signifies greater deviation from empirical data.

The sensitivity results across models with different DMRs (Fig. 3(A)) indicate that adding the OR interaction improved model sensitivity, whereas including the DF interaction resulted in reduced

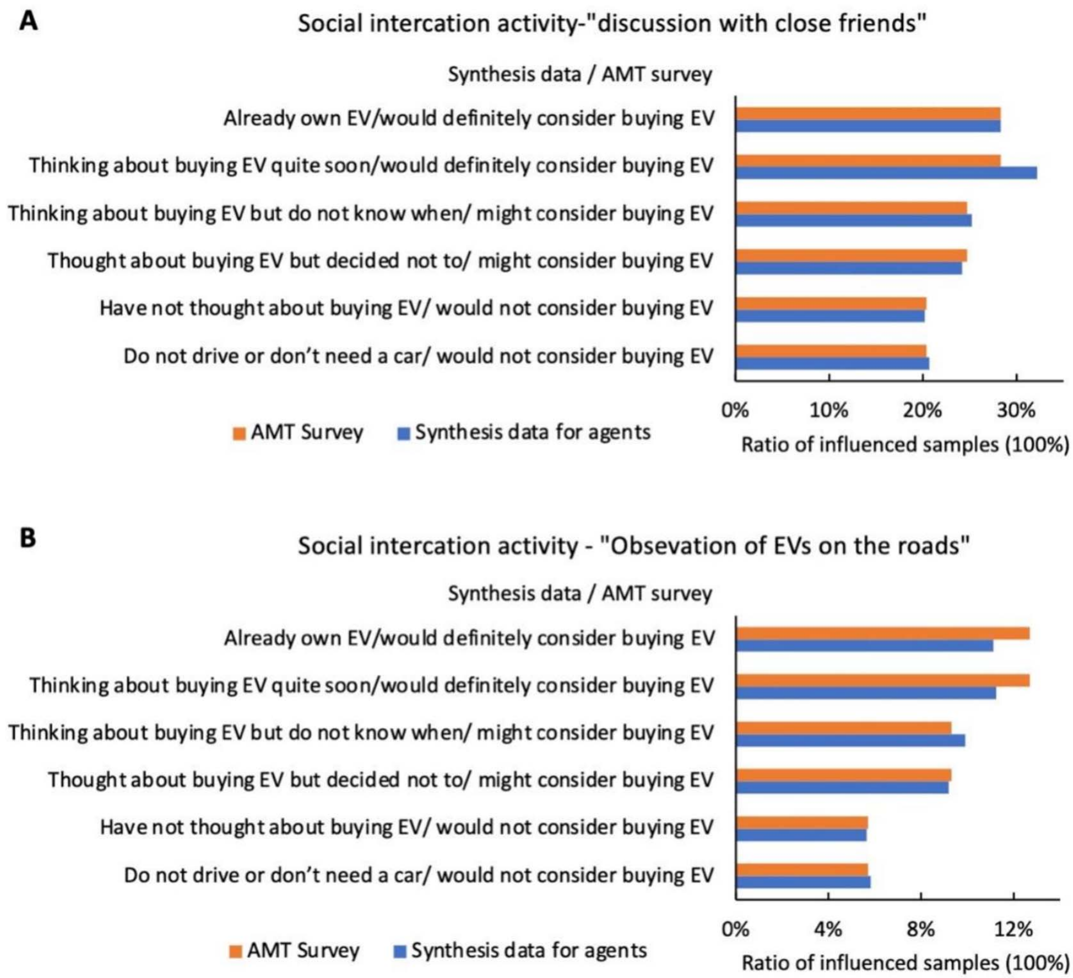


FIG. 2. Agents influenced by (A) 'discussions with close friends' and (B) 'observation of EVs on the roads' (AMT/Synthesized data).

sensitivity. The DMR_LR+OR rule improved the true positive rate, correctly simulating 71.05% (502) of positive cases as positive. Additionally, 28.95% (204) of positive cases were incorrectly labelled as negative, indicating a reduction in Type I errors compared to the DMR_LR rule, which lacks the social interaction feature. Regarding specificity (Fig. 3(B)), the DMR_LR+DF was superior, accurately simulating 70.1% (2,029) of negative cases as negative.

The overall model performance was assessed using the *G*-mean, which represents the geometric mean of sensitivity and specificity. Figure 3(C) shows that the DMR_LR+OR rule achieved a higher *G*-mean than the rule incorporating DF interaction, making DMR_LR+OR the preferred DMR. This rule captures individual attitudes with acceptable accuracy by combining the data-driven logistic regression function with the OR interaction, which reflects the provincial norm effect through 'observation of EVs on the roads'.

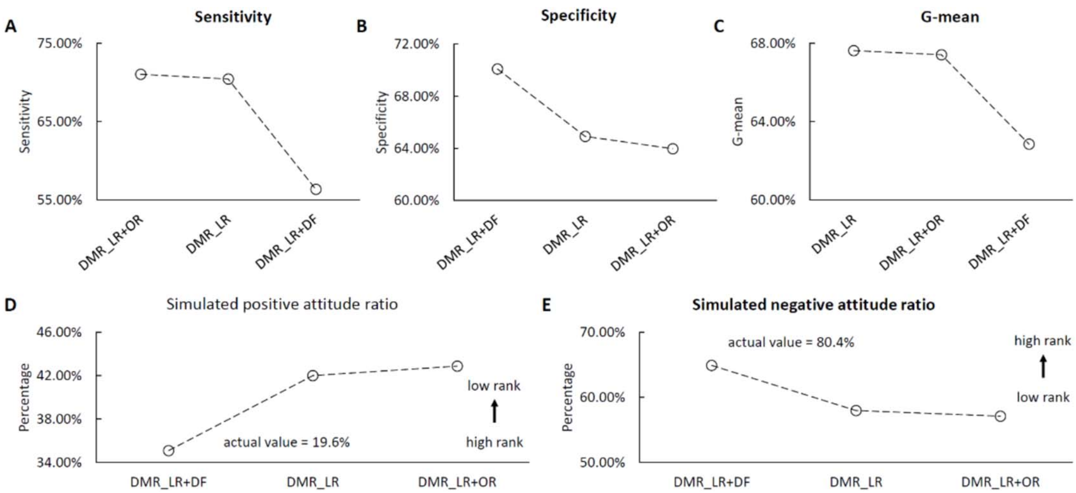


FIG. 3. Performance comparison: DMR_LR, DMR_LR+DF and DMR_LR+OR.

Figures 3(D and E) highlight how different decision rules affect population attitude simulations, revealing discrepancies linked to Type I and Type II errors (Nisbet *et al.*, 2009). A differences between individual and population simulations underscores the importance of micro-level validation in ABM to accurately capture individual heterogeneity. Previous models' inaccuracies may stem from insufficient micro-level data, sampling biases and uncalibrated decision rules. Validating decision rules at the micro level ensures that simulated attitudes align with actual attitudes, thereby reducing errors and enhancing ABM accuracy.

4.3. *SML-PV validator to evaluate the impact of social interaction parameters on agent-based model performances*

Social interaction parameters were initially adopted from the AMT data (Krupa *et al.*, 2014) and evaluated through parameter variation experiments to improve model performance. Specifically, OR interaction controlling parameters, ORprobability, were selected from a predefined value range to generate various DMRs for agents. Applying these rules within an agent-based model produced a dataset of agents' OR interaction features, which SML algorithms analyzed performance patterns and reduce parameter bias. This approach allowed for capturing agent heterogeneity in OR interaction features by optimizing model performance.

A pilot study was first conducted on six sets of ORprobability_i values to evaluate the model's performance under controlled conditions. Subsequently, a comprehensive parameter variation experiment using 15,625 sets of ORprobability_i values was conducted to reveal the relationship between OR interaction parameters and model performance. The SML-PV was used for validation, applying C5.0, CART and logistic regression algorithms.

Figure 4 visualizes the relationship between OR interaction parameters (ORprobability_1, ORprobability_4 and ORprobability_5) and model performance. The *G*-mean of DMR_LR+OR surpasses that of DMR_LR when $\text{ORprobability}_4 = 0.1395$, but was lower when $\text{ORprobability}_5 = 0.0285$, with $\text{ORprobability}_1 \in \{0.0000, 0.0635, 0.1270\}$ and other $\text{ORprobability}_i = 0$. These results suggest

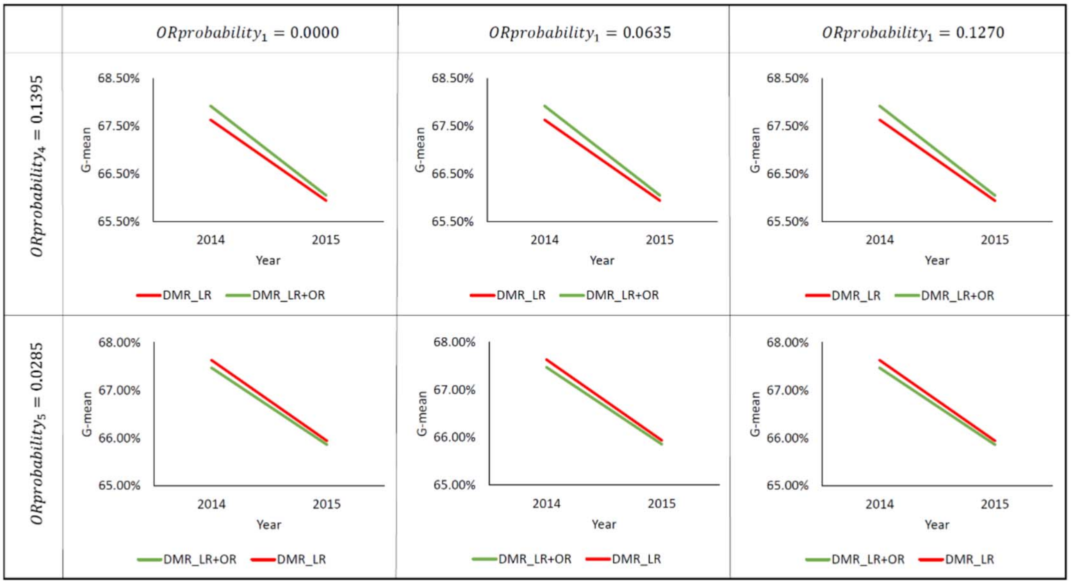


FIG. 4. Parameters variation experiment on ORprobability.

TABLE 4. *OR interaction controlling parameters*

OR interaction controlling parameters	Values
ORprobability_1	{0, 0.0635, 0.1270, 0.1905, 0.2540}
ORprobability_2	{0, 0.0635, 0.1270, 0.1905, 0.2540}
ORprobability_3	{0, 0.0465, 0.0930, 0.1395, 0.1860}
ORprobability_4	{0, 0.0465, 0.0930, 0.1395, 0.1860}
ORprobability_5	{0, 0.0285, 0.0570, 0.0855, 0.1140}
ORprobability_6	{0, 0.0285, 0.0570, 0.0855, 0.1140}

enhancing DMR_LR+OR *G*-mean by increasing OR interaction probability, especially for agents who considered buying EVs but refrained.

Comprehensive parameter experiments analyzed how adjustments to social interaction controlling parameters could improve agent-based model performance. Optimized parameters were validated for application in attitude classification datasets (e.g., ONS) lacking initially collected social interaction features.

The OR interaction controlling parameters were restricted to an explorable range to reduce iterations and runtime. This range (Table 4) is reasonable, aligning with values suggested by Bass (1969) and Rand & Rust (2011), with agent influence probability ranging from 0.026 to 0.6541 for potential adopters. This approach resulted in 15,625 ORprobability parameter combinations (ORprobability_1–ORprobability_6) used to evaluate model performance.

The *GmeanImprovement* metric was used to assess DMR performance. A value of 0 indicated that for ORprobability_i, the *G*-mean of the DMR_LR+OR was lower than or equal to that of DMR_LR,

TABLE 5. *Metrics for SML-PV in 5-data-point parameter variation experiment for* ORprobability *validation*

SML-PV candidates	**AUC (Mean, SD)**	**Sensitivity (Mean, SD)**	**Specificity (Mean, SD)**	***G*-mean (Mean, SD)**
ORprobability Logistics regression	(94.16%, 0.56%)	(88.59%, 2.8%)	(84.66%, 0.93%)	(86.60%, 5.3%)
C5.0	(99.46%, 0.11%)	(99.44%, 0.49%)	(96.94%, 0.49%)	(98.18%, 0.27%)
CART	(97.34%, 0.71%)	(95.63%, 1.89%)	(91.37%, 1.42%)	(93.47%, 0.88%)

TABLE 6. *DF interaction controlling parameters*

DF interaction controlling parameters	**Values in 5-data-points parameters variation experiment**
DFprobability_1	{0, 0.1415, 0.2830, 0.4245, 0.5660}
DFprobability_2	{0, 0.1415, 0.2830, 0.4245, 0.5660}
DFprobability_3	{0, 0.1235, 0.2470, 0.3705, 0.4940}
DFprobability_4	{0, 0.1235, 0.2470, 0.3705, 0.4940}
DFprobability_5	{0, 0.1020, 0.2040, 0.3060, 0.4080}
DFprobability_6	{0, 0.1020, 0.2040, 0.3060, 0.4080}

while a value of 1 indicated the opposite. The SML-PV classifier was constructed to recognize patterns in *GmeanImprovement*, visualize data, and determine suitable ORprobability parameter sets for OR social interaction formulation. SML-PV classifiers were trained and validated using the 2014 results (Section 3.2).

Table 5 presents evaluation metrics for SML-PV with ORprobability. The C5.0 algorithm outperforms others, achieving a mean AUC of 99.46%, sensitivity of 99.44%, specificity of 96.94% and G-mean of 98.18%. Testing the SML-PV with C5.0 on the 2015 simulation yielded an AUC of 97.80%, sensitivity of 99.09%, specificity of 89.71% and G-mean of 94.28%, demonstrating generalizability. This approach effectively recognized the relationship between OR social interaction controlling parameters and DMR_LR+OR performance, largely attributed to simplified OR interaction patterns and *GmeanImprovement*. Future research could explore more complex relationships with enhanced computing power. Appendix A presents the first five layers of the SML-PV with C5.0 for OR interaction parameters, showing consistency with AMT empirical data (Section 4.2).

DF interaction parameters were also examined using a 5-data-point parameter variation experiment (Table 6), validated with SML-PV. The results, presented in Table 7, indicate that C5.0 performed best across all metrics. Testing SML-PV with C5.0 on the 2015 parameter variation experiment yielded an AUC of 99.80%, sensitivity of 97.86%, specificity of 96.34% and G-mean of 97.10%, showing strong generalizability. Appendix B visualizing the first five layers of the SML-PV with C5.0 for DF interaction parameters.

TABLE 7. *Metrics for SML-PV in 5-data-point parameters variation experiment for* DFprobability *validation*

SML-PV candidates	AUC (Mean, SD)	Sensitivity (Mean, SD)	Specificity (Mean, SD)	*G*-mean (Mean, SD)
DFprobability				
Logistic regression	(95.06%, 0.65%)	(94.47%, 1.08%)	(85.48%, 0.66%)	(89.86%, 0.72%)
C5.0	(99.50%, 0.35%)	(98.15%, 1.32%)	(97.72%, 1.04%)	(97.93%, 0.68%)
CART	(97.34%, 0.71%)	(95.63%, 1.89%)	(91.37%, 1.42%)	(93.47%, 0.88%)

The OR and DF interaction values with the highest *G*-mean were selected to parameterize the final optimized DMR_LR+OR and DMR_LR+DF, respectively. The highest *G*-mean value signifies the closest alignment with empirical data for the predicted positive attitude ratio.

In summary, SML-PV classifiers determine optimized social interaction parameters, enhancing their efficacy in simulating consumer attitudes towards EV purchasing via agent-based models. When DF and OR interaction controlling parameters are empirically accessible, the social interaction components of the ABM can be further refined for case-specific attitude classifications. This data-driven approach minimizes biases in synthesized social interaction features, facilitating the optimization of DMRs to closely align with ONS respondents' attitudes towards EV adoption. As a result, this model provides a robust, adaptable tool for accurately capturing consumer attitudes in dynamic, socially interactive environments.

4.4. *Agent-based model to simulate consumer attitudes*

The optimized DMRs DMR_LR+OR and DMR_LR+DF were applied to simulate ONS consumer attitudes, as illustrated in Fig. 5. Figure 5(A) shows the sensitivity of each model for both 2014 and 2015, with the optimized DMR_LR+OR outperforming the others. Figure 5(B) represents specificity, where the optimized DMR_LR+DF performs better, while the optimized DMR_LR+OR has lower specificity. Figure 5(C) demonstrates *G*-mean, where the optimized DMR_LR+OR again leads in both years. When simulating consumer attitudes using 2014 data, the optimized DMR_LR+OR model achieved a sensitivity of 71.54%, specificity of 64.92% and a *G*-mean of 68.15%, significantly outperforming the DMR_LR rule. A similar performance pattern was observed with the 2015 data, indicating consistent model accuracy across different datasets.

Furthermore, the optimized DMR_LR+OR outperformed the rule incorporating the OR interaction synthesized from the AMT data (as described in Section 4.2). These results confirm that the optimized DMR effectively captures the nuances of consumer attitudes towards EV adoption in the context of ONS data, offering a robust tool for simulating and analyzing consumer behaviour in socially interactive settings.

5. Discussion and conclusion

This study presents a novel integrated model that combines ABM and SML methods to explore the relationship between consumer attitudes and social interactions in the adoption of EVs. The integrated model identifies key consumer features from empirical data and synthesizes missing or 'unseen' social

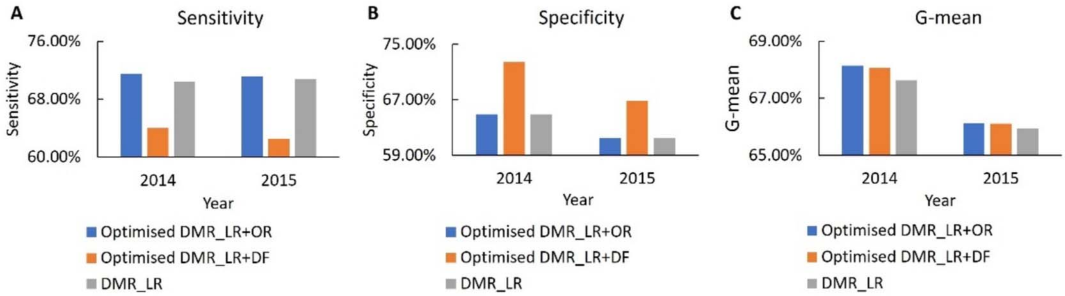


FIG. 5. Performance of optimized DMR_LR+OR, optimized DMR_LR+DF and DMR_LR.

interaction data, thereby accounting for consumer heterogeneity in attitude simulations. Through micro-level model validation and optimization, the method leverages ONS survey data from 2014 and 2015 to simulate consumer attitudes towards EV adoption in Great Britain.

Our results underscore the critical influence of social interactions in shaping consumer attitudes during the early stage of EV diffusion, a period marked by low adoption rates and limited purchasing behaviour data. This work offers a template for applying attitude-intention-behaviour theories in a comprehensive, bottom-up approach to consumer attitude analysis.

This research introduces a modularized model that not only identifies key consumer features and classifies attitudes, but also synthesizes social interaction data to address information gaps. For instance, the model augments missing data on social influence effects by integrating insights from relevant theories and additional sources like AMT data. Furthermore, micro-validation and optimization of consumers' DMRs ensure that the model accurately reflects consumer heterogeneity in social interactions and their impacts on attitude formation.

Key findings reveal five influential factors driving EV adoption attitudes: environmental concerns, importance of technology, battery range limitations, car ownership and education level. Incorporating social interactions into the DMRs, such as 'discussions with close friends' and 'observation of EVs on roads', significantly improved the model's performance in attitude classification.

Our findings suggest that marketers should develop tailored strategies targeting specific consumer segments to encourage EV adoption. This recommendation is based on the observation that consumers with similar socio-demographic profiles exhibit similar attitudes towards EV adoption. These strategies should be periodically updated to reflect different stages of the adoption process. Additionally, advancing battery technology and capital investments are essential to enhancing EV battery range, which remains a significant consumer concern.

Policy implications from this work suggest that policymakers should prioritize public awareness campaigns on environmental protection in Great Britain, as our finding indicate that individuals with strong environmental concerns are more likely to adopt positive attitudes towards EVs. Effective measures may include public awareness campaigns through social media, TV and print advertisements, environmental education in schools and community engagement initiatives.

Moreover, our study provides empirical recommendations for fostering positive attitudes towards EV adoption by promoting social activities and increasing EV visibility. Findings indicate that positive social interactions with close friends and greater visibility of EVs in residential areas contribute to favourable consumer attitudes. Suggested measures include EV demonstrations at community events, EV-friendly festivals and exhibitions, expanded charging facilities in prominent locations, EV ridesharing and taxi services and collaborations with influencers and environmental advocates.

The findings of this study align with Social Influence Theory and extend the TPB by demonstrating the role of social interactions in sharping attitudes towards EV adoption. We propose enhancements to the theoretical framework, highlighting the mediating role of consumer attitudes in the relationship between subjective norms and behavioural intentions. By validating DMRs at the micro level, our approach mitigates potential evaluative biases in agent-based models, facilitating the application of Diffusion of Innovations Theory to better understand attitudes towards innovation adoption.

Compared to the work of Liu & Xiao (2018), which used a top-down system dynamics approach to predict EV adoption, our model enables tracking individual consumer behaviours. Given the availability of empirical data for model validation (data from 2014 and 2015 in this study), our model assesses how accurately consumers' attitudes were predicted across both years using evaluation metrics, such as sensitivity, specificity and *G*-mean. Unlike top-down approaches, our model offers more granular insights that align closely with real-world data, demonstrating its effectiveness in capturing the complexities of decision-making in EV adoption.

Potential extensions to this model include the dynamic incorporation of various forms of social influences and periodic updates to DMRs with new empirical data. Future research could also integrate intentional and behavioural components for a complete decision-making process, making the model applicable beyond EV adoption studies.

The study has several limitations. First, a more detailed categorization of attitudes related to the data would enhance accuracy. In addition, the Great Britain-specific data may limit the direct application of our findings to other regions or countries with distinct cultural and regulatory environments. For example, our research identifies that environmental concerns positively influence EV adoption among UK consumers, a result consistent with findings on Chinese consumers (Wu *et al.*, 2019). However, Qiao & Dowell (2022) found that environmental concerns play a less significant role in the adoption of environmentally friendly products like Tesla among US consumers, who prioritize vehicle performance. This variation may reflect cultural differences in the impact of environmental concerns on EV adoption. To improve generalizability, further studies should consider populations from diverse countries and regions. Nevertheless, our integrated ABM and SML approach can be directly applied to data from the other regions to investigate public attitudes towards EV adoption. Furthermore, while the model may not capture all dimensions, it establishes a foundation for more comprehensive research on consumer attitudes in the context of innovation adoption.

Acknowledgements

The authors would like to acknowledge valuable research assistance support from Dr Paul Nieuwenhuis at Cardiff Business School. They also thank Prof. Peter Morgan (Cardiff Business School) for useful opinions on supervised machine learning modelling, Prof. Peter Lugtig (Utrecht University) for help on survey data preparation and technical support from Ms. Tatiana Gomzina (AnyLogic Technical Support Team) on agent-based modelling. They would also like to thank Prof. David Barrow (Cardiff School of Engineering, Cardiff University) for useful conversations and comments on the model and results.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Conflict of interest

None declared.

Data availability

The data underlying this article were provided by Office for National Statistics, Social Survey Division and UK Data Service under licence/by permission. Data will be shared on request to the corresponding author with permission of Office for National Statistics, Social Survey Division and UK Data Service.

References

Abbasi, S., Mousavi, S. S., Farbod, E., Yousefi Sorkhi, M. & Parvin, M. (2024) Hybrid data mining and data-driven algorithms for a green logistics transportation network in the post-COVID era: A case study in the USA. *Syst. Soft. Comput.*, **6**, 200156. https://doi.org/10.1016/j.sasc.2024.200156.

Aboutaleb, Y. M., Danaf, M., Xie, Y. & Ben-Akiva, M. (2021) *Discrete Choice Analysis with Machine Learning Capabilities [Preprint]* Available at: https://arxiv.org/abs/2101.10261. Date accessed 24 August 2021.

Agliari, E., Burioni, R., Cassi, D. & Maria Neri, F. (2010) Word-of-mouth and dynamical inhomogeneous markets: An efficiency measure and optimal sampling policies for the pre-launch stage. *IMA J. Manag. Math.*, **21**, 67–83.

Ajzen, I. (1985) From intentions to actions: A theory of planned behaviour. *Action Control: From Cognition to Behaviour* (J. Kuhi & J. Bechkmann eds.). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 11–39.

An, L. (2012) Modelling human decisions in coupled human and natural systems: Review of agent-based models. *Ecol. Model.*, **229**, 25–36.

Arts, J. W. C., Frambach, R. T. & Bijmolt, T. H. A. (2011) Generalizations on consumer innovation adoption: A meta-analysis on drivers of intention and behaviour. *Int. J. Res. Mark.*, **28**, 134–144.

Axsen, J. & Kurani, K. S. (2014) Social influence and proenvironmental behaviour: The reflexive layers of influence framework. *Environ. Plann. B Plann. Des.*, **41**, 847–862.

Baltas, I. & Yannacopoulos, A. N. (2019) Portfolio management in a stochastic factor model under the existence of private information. *IMA J. Manag. Math.*, **30**, 77–103.

Bananuka, J., Kasera, M., Najjemba, G. M., Musimenta, D., Ssekiziyivu, B. & Kimuli, S. N. L. (2019) Attitude: Mediator of subjective norm, religiosity and intention to adopt Islamic banking. *J. Islam. Mark.*, **11**, 81–96.

Barsyte, J. & Fennis, B. M. (2023) When innovation backfires: Preference for predictability moderates the spillover of functional food ambivalence to the entire parent category. *J. Bus. Res.*, **158**, 113670. https://doi.org/10.1016/j.jbusres.2023.113670.

Barth, M., Judgert, P. & Fritsche, I. (2016) Still underdetected – Social norms and collective efficacy predict the acceptance of electric vehicles in Germany. *Transp. Res. F Traffic Psychol. Behav.*, **37**, 64–77.

Bass, F. M. (1969) A new product growth for model consumer durables. *Manag. Sci.*, **15**, 215–227.

Bell, D. & Mgbemena, C. (2018) Data-driven agent-based exploration of customer behaviour. *Simulation*, **94**, 195–212.

Bennett, R. & Vijaygopal, R. (2018) Consumer attitudes towards electric vehicles: Effect of product user steretypes and self-image congruence. *Eur. J. Mark.*, **52**, 499–527.

Berkeley, N., Javis, D. & Jones, A. (2018) Analysing the take up of battery electric vehicles: An investigation of barriers amongst drivers in the UK. *Transp. Res. D*, **63**, 466–481.

Brathwaite, T., Vij, A. & Walker, J. L. (2017) *Machine Learning Meets Microeconomics: The Case of Decision Trees and Discrete Choice ArXiv:1711.04826v1, Cornell University*. Available at: https://arxiv.org/abs/1712.04429. Date accessed 7 July 2021.

Carley, S., Krause, R. M., Lane, B. W. & Graham, J. D. (2013) Intent to purchase a plug-in electric vehicle: A survey of early impressions in large US cites. *Transp. Res. D*, **18**, 39–45.

Castro-Santa, J., Drews, S. & van den Bergh, J. (2023) Nudging low-carbon consumption through advertising and social norms. *J. Behav. Exp. Econ.*, **102**, 101956. https://doi.org/10.1016/j.socec.2022.101956.

Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. (2002) SMOTE: Synthetic minority over-sampling technique. *J. Artificial Intelligence Res.*, **16**, 321–357.

Cialdini, R. B., Reno, R. R. & Kallgren, C. A. (1990) A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *J. Pers. Soc. Psychol.*, **58**, 1015–1026.

Clark, R. D. (2003) Boosted leave-many-out cross-validation: The effect of training and test set. *J. Comput.-Aided Mol. Des.*, **17**, 265–275.

DfT (Department for Transport) (2016) *Public Attitudes towards Electric Vehicles: 2016 (Revised)* Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/551446/electric-vehicles-survey-2016.pdf. Date accessed 20 July 2021.

DfT (Department for Transport) (2020) *Government Takes Historic Step towards Net-Zero with End of Sale of New Petrol and Diesel Cars by 2030* Available at: https://www.gov.uk/government/news/government-takes-historic-step-towards-net-zero-with-end-of-sale-of-new-petrol-and-diesel-cars-by-2030. Date accessed 6 September 2021.

Eppstein, M. J., Grover, D. K., Marshall, J. S. & Rizzo, D. M. (2011) An agent-based model to study market penetration of plug-in hybrid electric vehicles. *Energy Policy*, **39**, 3789–3802.

Eppstein, M. J., Rizzo, D. M., Lee, B. H. Y., Krupa, J. S. & Manukyan, N. (2015) Using national survey respondents as consumers in an agent-based model of plug-in hybrid vehicle adoption. *IEEE Access*, **3**, 457–468.

Fishbein, M. (1979) A theory of reasoned action: Some applications and implications. *Nebr. Symp. Motiv.*, **27**, 65–116.

Fishbein, M. & Ajzen, I. (1975) *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research*. Boston: Addison-Wesley Publishing Company.

Flynn, T., Tian, Y., Masnick, K., McDonnell, G., Huynh, E., Mair, A. & Osgood, N. (2014) Discrete choice, agent based and system dynamics simulation of heath profession career paths. *Proceedings of the 2014 Winter Simulation Conference, Savanah, Georgia, 7–10 December 2014* (A. Tolk et al. eds.). Piscataway, NJ: IEEE, pp. 1700–1711.

Foramitti, J., Savin, I. & van den Bergh, J. C. J. M. (2024) How carbon pricing affects multiple human needs: An agent-based model analysis. *Ecol. Econ.*, **217**, 108070. https://doi.org/10.1016/j.ecolecon.2023.108070.

Fry, A., Reley, T. & Thring, R. (2018) The influence of knowledge and persuasion on the decision to adopt or reject alternative fuel vehicles. *Sustainability*, **10**, 2997.

Furtado, B. A. (2020) Machine learning simulates agent based model. *CoMESE Network, Computational Model Library, 7 March 2018* Available at: https://www.comses.net/codebases/59d377e4-afa3-4937-a6fc-30cc50c1ac8f/releases/1.0.0/. Date accessed 19 May 2020.

Gallagher, S., Richardson, L. F., Ventura, S. L. & Eddy, W. F. (2018) SPEW: Synthetic populations and ecosystems of the world. *J. Comput. Graph. Statist.*, **27**, 773–784.

Garcia, R. (2005) Use of agent-based modelling in innovation/new product development research. *Prod. Dev. Manag. Assoc.*, **22**, 380–398.

Goldenberg, J., Libai, B., Solomon, S., Jan, N. & Stauffer, D. (2000) Marketing percolation. *Physica A Stat. Mech. Appl.*, **284**, 335–347.

Grindrod, P. & Higham, D. J. (2012) Models for evolving networks: With applications in telecommunication and online activities. *IMA J. Manag. Math.*, **23**, 1–15.

Gürcan, Ö., Dikenelli, O. & Bernon, C. (2013) A generic testing framework for agent-based simulation models. *J. Simul.*, **7**, 183–201.

Guseo, R. & Mortarino, C. (2014) Within-brand and cross-brand word-of-mouth for sequential multi-innovation diffusions. *IMA J. Manag. Math.*, **25**, 287–311.

Gómez Vilchez, J. J. & Jochem, P. (2019) Simulating vehicle fleet composition: A review of system dynamics models. *Renew. Sustain. Energy Rev.*, **115**, 109367.

Gómez Vilchez, J. J., Jochem, P. & Fichtner, W. (2013) EV market development pathways – An application of system dynamics for policy simulation. *World Electr. Veh. J.*, **6**, 1030–1038.

Hamill, L. & Gilbert, N. (2009) Social circles: A simple structure for agent-based social network models. *J. Artif. Soc. Soc. Simul.*, **12**, 1–23.

Hamill, L. & Gilbert, N. (2016) *Agent-Based Modelling in Economics*. West Sussex: John Wiley & Sons, Ltd.
Hassouna, M. B. (2012) *Agent Based Modelling and Simulation: An Examination of Customer Retention in the UK Mobile Market* PhD Thesis,. Uxbridge, West London: Brunel University.
He, F., Wu, D., Yin, Y. & Guan, Y. (2013) Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transp. Res. B Methodol.*, **47**, 87–101.
He, F., Yin, Y. F. & Lawphonganich, S. (2014) Network equilibrium models with battery electric vehicles. *Transp. Res. B Methodol.*, **67**, 306–319.
Hoffmann, J. P. (2016) *Regression Models for Categorical, Count, and Related Variables: An Applied Approach*. Berkeley, CA: University of California Press.
IBM (2018) *IBM SPSS Modeler 18.2 Algorithms Guide*, Armonk, NY: IBM Corporation, pp. 157–166.
IEA (2021) *Global EV Outlook 2021: Accelerating Ambitions despite the Pandemic*. Paris, France: IEA. Available at: https://iea.blob.core.windows.net/assets/ed5f4484-f556-4110-8c5c-4ede8bcba637/GlobalEVOutlook2021.pdf. Date accessed 24 September 2021.
Iftekhar, M. S., Hailu, A. & Lindner, R. K. (2011) Item price information feedback in multiple unit combinatorial auctions: Design issues. *IMA J. Manag. Math.*, **22**, 271–289.
Jäger, G. (2019) Replacing rules by neural networks: A framework for agent-based modelling. *Big Data Cogn. Comput.*, **3**, 51–63.
Japkowicz, N. (2013) Chapter 8: Assessment metrics for imbalanced learning. *Imbalanced Learning: Foundations, Algorithms, and Applications* (H. He & Y. Q. Ma eds.). New Jersey, US: Wiley, pp. 187–206.
Kelman, H. C. (1958) Compliance, identification, and internalization: Three processes of attitude change. *J. Conflict Resolut.*, **2**, 51–60.
Kiesling, E., Günther, M., Stummer, C. & Wakolbinger, L. M. (2012) Agent-based simulation of innovation diffusion: A review. *Cent. Eur. J. Oper. Res.*, **20**, 183–230.
Kim, T., Hong, J.-S. & Lee, H. (2016) Predicting when the mass market starts to develop: The dual market model with delayed entry. *IMA J. Manag. Math.*, **27**, 381–396.
Kim, M., Park, J.-H., Kim, K. & Park, B. (2020) Identifying factors influencing the slow market diffusion of electric vehicles in Korea. *Transportation*, **47**, 663–688.
Kokkinaki, F. & Lunt, P. (1997) The relationship between involvement, attitude accessibility and attitude-behaviour consistency. *Br. J. Soc. Psychol.*, **36**, 497–509.
Kotthoff, F. & Hamacher, T. (2022) Calibrating agent-based models of innovation diffusion with gradients. *J. Artif. Soc. Soc. Simul.*, **25**, 4.
Krupa, J. S., Rizzo, D. M., Eppstein, M. J., Lanute, D. B., Gaalema, D. E., Lakkaraju, K. & Warrender, C. (2014) Analysis of a consumer survey on plug-in hybrid electric vehicles. *Transp. Res. A*, **64**, 14–31.
Lamperti, F., Roventini, A. & Sani, A. (2017) Agent-based model calibration using machine learning surrogates. *J. Econom. Dynam. Control*, **90**, 366–389.
Lane, B. & Potter, S. (2007) The adoption of cleaner vehicles in the UK: Exploring the consumer attitude-action gap. *J. Clean. Prod.*, **15**, 1085–1092.
Liu, D. & Xiao, B. (2018) Exploring the development of electric vehicles under policy incentives: A scenario-based system dynamics model. *Energy Policy*, **120**, 8–23.
Liu, J., Nie, J. & Yuan, H. (2023) Electric vehicle manufacturers' decisions on investing in carbon-reduction technology under government subsidy: A Cournot game model. *IMA J. Manag. Math.*, **34**, 71–100.
Luengo, J., Fermández, A., García, S. & Herrera, F. (2010) Addressing data complexity for imbalanced data sets: Analysis of SMOTE-based oversampling and evolutionary undersampling. *Soft Comput.*, **15**, 1909–1936.
Lui, M., Sin, K. F., Yang, L., Forlin, C. & Ho, K. C. (2015) Knowledge and perceived social norm predict parents' attitudes towards inclusive education. *Int. J. Incl. Educ.*, **19**, 1052–1067.
Mahajan, V. & Muller, E. (1979) Innovation diffusion and new product growth models in marketing. *J. Mark.*, **43**, 55–68.
Mairesse, O., Macharis, C., Lebeau, K. & Turcksin, L. (2012) Understanding the attitude-action gap: Functional integration of environmental aspects in car purchase intentions. *Psicológica*, **33**, 547–574.

Mao, A., Huang, E., Gan, H., Parkes, R. S. V., Xu, W. & Liu, K. (2021) Cross-modality interaction network for equine activity recognition using imbalanced multi-modal data. *Sensors*, **21**, 5818.

Martin, D. M., Väistö, T. & Malhotra, N. K. (2016) Reducing the attitude-behaviour gap in sustainable consumption: A theoretical proposition and the American electric vehicle market. *Mark. Sustain. Soc.*, **13**, 193–213.

McFadden, D. (1981) Econometric models of probabilistic choice. *Structural Analysis of Discrete Data with Econometric Applications* (D. McFadden & C. F. Manski eds.). Cambridge: MIT, pp. 198–272.

Moulaei, K., et al. (2024) Explainable artificial intelligence (XAI) for predicting the need for intubation in methanol-poisoned patients: A study comparing deep and machine learning models. *Sci. Rep.*, **14**, 15751.

Nie, Y. M., Ghamami, M., Zockaie, A. & Xiao, F. (2016) Optimization of incentive polices for plug-in electric vehicles. *Transp. Res. B Methodol.*, **84**, 103–123.

Nisbet, R., Miner, G. & Yale, K. (2009) *Handbook of Statistical Analysis and Data Mining Applications*, 2nd edn. London, UK: Elsevier.

Onat, N. C., Kucukvar, M. & Tatari, O. (2014) Conventional, hybrid, plug-in hybrid or electric vehicles? State-based comparative carbon and energy footprint analysis in the United States. *Appl. Energy*, **150**, 36–49.

ONS (Office for National Statistics), Social Survey Division (2015) *Opinions and Lifestyle Survey, Electric Vehicles Module, February 2014 and February 2015*. Available at: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7571&type=Data%20catalogue#!/details. Date accessed 1 May 2018.

Ozdemir, S. (2016) Chapter 12: Beyond the essentials. *Principles of Data Science: Learn the Techniques and Match your Need to Start Making Sense of your Data* (S. Vernekar et al. eds.). Birmingham, UK: Packt Publishing, pp. 289–324.

Peattie, K. (2010) Green consumption: Behavior and norms. *Annu. Rev. Environ. Resour.*, **35**, 195–228.

Pertiwi, A. G., Bachtiar, N., Kusumaningrum, R., Waspada, I. & Wibowo, A. (2020) Comparison of performance of K-nearest neighbour algorithm using smote and k-nearest neighbour algorithm without smote in diagnosis of diabetes disease in balanced data. *J. Phys. Conf. Ser.*, **1524**, 012048. https://doi.org/10.1088/1742-6596/1524/1/012048.

Peysakhovich, A. & Naecker, J. (2017) Using methods from machine learning to evaluate behavioral models of choice under risk and ambiguity. *J. Econ. Behav. Organ.*, **133**, 373–384.

Prislin, R. & Wood, W. (2005) Social influence in attitude and attitude change. *The Handbook of Attitude* (D. Albarracín, B. T. Johnson & M. P. Zanna eds.). Mahwah, NJ: Lawrence Erlbaum Associates Publishers, pp. 671–705.

Qiao, K. & Dowell, G. (2022) Environmental concerns, income inequality, and purchase of environmentally-friendly products: A longitudinal study of U.S. counties (2010-2017). *Res. Policy*, **51**, 104443.

Rand, W. & Rust, R. T. (2011) Agent-based modelling in marketing: Guidelines for rigor. *Int. J. Res. Mark.*, **28**, 181–193.

Ranganathan, P., Pramesh, C. S. & Aggarwal, R. (2017) Common pitfalls in statistical analysis: Logistic regression. *Perspect. Clin. Res.*, **8**, 148–151.

Ravaioli, G., Domingos, T., Teixeira, F. M. & R. (2023) Data-driven agent-based modelling of incentives for carbon sequestration: The case of sown biodiverse pastures in Portugal. *J. Environ. Manag.*, **338**, 117834. https://doi.org/10.1016/j.jenvman.2023.117834.

Robinson, S. (2014) *Simulation: The Practice of Model Development and Use*, 2nd edn. London: Macmillan Education UK.

Rogers, E. M. (1976) New product adoption and diffusion. *J. Consum. Res.*, **2**, 290–301.

Rogers, E. M. (2003) *Diffusion of Innovations. 4th Version*. New York: The Three Press.

Sarle, W. S. (2002) *Subject: What Are the Population, Sample, Training Set, Design Set, Validation Set, and Test Set? Nerual Network FAQ, part 1 of 7: Introduction*. Cary, NC: Warren S. Sarle. Available at: http://francky.me/aifaq/FAQ-comp.ai.neural-net.pdf. Date accessed 28 July 2021.

Schmalfuß, F., Mühl, K. & Krems, J. F. (2017) Direct experience with battery electric vehicles (BEVs) matters when evaluating vehicle attributes, attitude and purchase intention. *Transp. Res. F Traffic Psychol. Behav.*, **46**, 47–69.

Shim, D., Shin, J. & Kwak, S. Y. (2018) Modelling the consumer decision-making process to identify the key drivers and bottlenecks in the adoption of environmentally friendly products. *Bus. Strateg. Environ.*, **27**, 1409–1421.

Sierzchula, W., Bakker, S., Maat, K. & van Wee, B. (2014) The influence of financial incentives and other socio-economic factors on electric vehicle adoption. *Energy Policy*, **68**, 183–194.

Song, Y. Y. & Lu, Y. (2015) Decision tree methods: Applications for classification and prediction. *Shanghai Arch. Psychiatry*, **27**, 130–135.

Song, R. Q. & Potoglou, D. (2020) Are existing battery electric vehicles adoption studies able to inform policy? A review for policy makers. *Sustainability*, **12**, 64–94.

van Staden, H. E., Yee, H. & Boute, R. N. (2024) Evaluating the cost-emissions trade-offs of a modal shift in intermodal and synchromodal transportation planning. *IMA J. Manag. Math.* (in press). https://doi.org/10.1093/imaman/dpae010.

UK Data Service (2014) University of Essex, Institute for Social and Economic Research, University of Manchester, Cathie Marsh Centre for Census and Survey Research. *Understanding Society, Wave 3, 2011–2012* [data collection]. Colchester, UK: UK Data Service (University of Essex). Available at: https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=7549. Date accessed 24 March 2021.

Vaus, D. (2002) *Analysing Social Science Data: 50 Key Problems in Data Analysis*. London: SAGE.

Wan, C., Shen, G. Q. & Choi, S. (2017) Experimental and instrumental attitudes: Interaction effect of attitude and subjective norm on recycling intention. *J. Environ. Psychol.*, **50**, 69–79.

Wardle, J., Blythe, P., Gibbon, J., Herron, C. & Hübner, Y. (2015) The provision of public recharging infrastructure for electric vehicles in the UK – Is there a business case? *World Electr. Veh.*, **7**, 546–557.

Widiarta, H., Viswanathan, S. & Piplani, R. (2008) Forecasting item-level demands: An analytical evaluation of top–down versus bottom–up forecasting in a production-planning framework. *IMA J. Manag. Math.*, **19**, 207–218.

Wolinetz, M. & Axsen, J. (2017) How policy can build the plug-in electric vehicle market: Insights from the respondent-based preference and constraints (REPAC) model. *Technol. Forecast. Soc. Change*, **117**, 238–250.

Wu, J., Liao, H., Wang, J.-W. & Chen, T. (2019) The role of environmental concern in the public acceptance of autonomous electric vehicles: A survey from China. *Transp. Res. F Traffic Psychol. Behav.*, **60**, 37–46.

Xing, R. X., Zhuang, M. M. & Li, B. (2024) For-profit or for-CSR: Decisions of power batteries recovery in the supply chain. *IMA J. Manag. Math.* (in press).

Zhang, H. & Vorobeychik, Y. (2017) Empirically grounded agent-based models of innovation diffusion: A critical review. *Artif. Intell. Rev.*, **52**, 707–741.

Zhang, H., Vorobeychik, Y., Letchford, J. & Lakkaraju, K. (2016) Data-driven agent-based modelling, with an application to rooftop solar adoption. *Auton. Agent. Multi-Ag. Syst.*, **30**, 1023–1049.

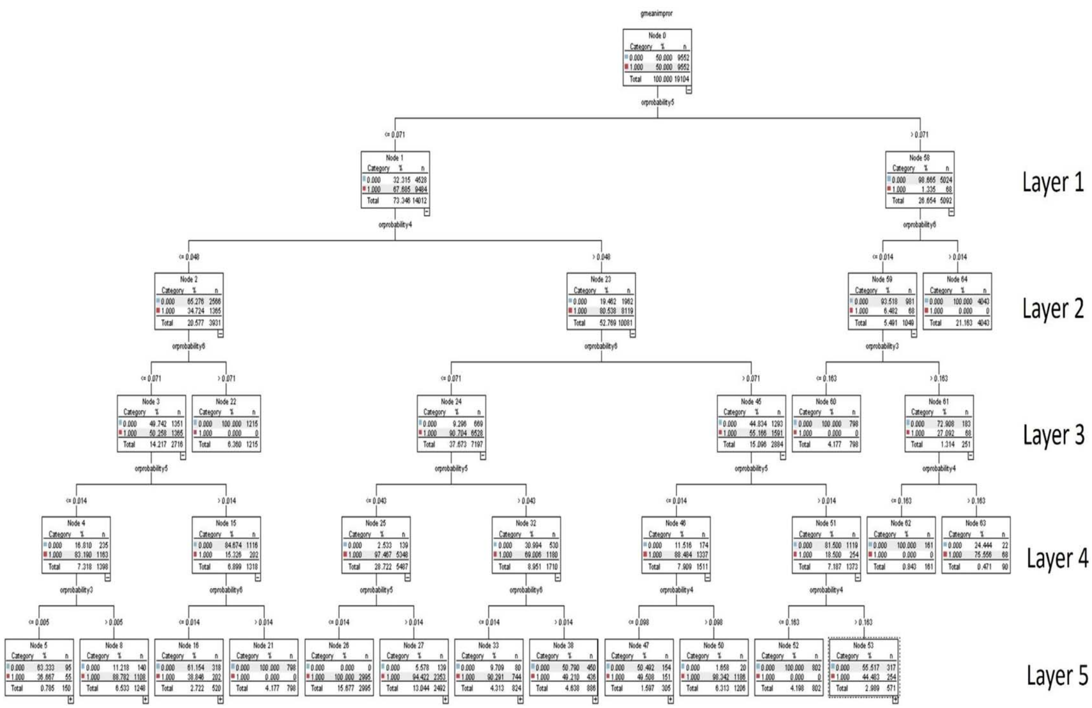
Zhou, Y. & Lund, P. D. (2023) Peer-to-peer energy sharing and trading of renewable energy in smart communities — trading pricing models, decision-making and agent-based collaboration. *Renew. Energy*, **207**, 177–193.

Appendix A. Comparison of the demographic distribution between the ONS and the AMT respondent populations

Demographic	ONS Survey (%)		AMT Survey (%)
	2014	2015	
Age			
16–24	8.8	6.9	29.5
26–44	28.4	30.2	52.2
45–54	17.7	15.0	12.1
55–64	16.9	18.6	5.1
65 and over	28.2	29.4	1.1
Gender			
Female	51.8	53.0	61.5
Male	48.2	47.0	38.5
Education level			
Degree or equivalent	23.9	25.5	43.0
Below degree level	43.9	43.8	11.0
Other qualifications	11.9	10.3	30.4
None	20.3	20.5	15.6
Income			
Lower than median income before tax	70.9	67.1	56.7
Equal or higher than median income before tax	29.1	32.9	43.3

The table compares the demographic distribution of the respondent populations from the ONS and the AMT surveys, including age, gender, education level, and income. From the table, we observe that the ONS survey respondents are similar to the AMT survey respondents in that (1) there are more female respondents than male and (2) a higher percentage of respondents have incomes below the median income before tax, compared to those with incomes at or above the median. However, AMT survey respondents tend to be younger and more highly educated. In this work, we assumed initial values on social interaction features of ONS agents based on ATM survey findings. These initial values are applied for initial model development and testing and have been validated in parameter variation experiments. As a result, the difference between AMT and ONS sample characteristics do not influence the experimental results.

Appendix B. The first five layers of SML-PV C5.0 based on 5-data-point parameters variation experiment on ORprobability



Appendix C. The first five layers of SML-PV C5.0 constructed based on 5-data-point parameters variation experiment on DFprobability