scientific reports

OPEN

Participatory inverse modelling for behavioural insights into Dutch EV charging dynamics

Lukas Schubotz[1]✉, Emile Chappin[1,3], Geeske Scholz[2,3] & Gerdien de Vries[1,3]

This paper presents a general epistemic workflow for behavioural inference in agent-based models under structural-parametric uncertainty by combining participatory and inverse modelling. This replicable computational pipeline enhances model realism, results explainability, and interpretability by integrating domain expertise. We apply our method to study Dutch EV charging behaviour. Using behavioural domain expertise, we co-design experiments and interpret alternative behavioural dynamics evaluated through inverse modelling in a structural-parametric pipeline. Public EV charging demand over a week in Den Haag is matched by simulating demand in an existing model of EV charging behaviour. Our method reveals that EV charging is largely driven by range anxiety or a lack of consideration for the availability of excess energy in the grid. While our behavioural findings are case-specific, the workflow is applicable to any agent-based model with ambiguous behavioural mechanisms and suitable for feeding back into theory, scrutinising formalisation, identifying areas for model improvement, and raising considerations for policy.

Keywords Inverse modelling, Participatory modelling, Behavioural theory, Structural calibration, Dutch energy transition, EV charging behaviour, Sense-making

Electric vehicles (EVs) are central to the Dutch strategy for reducing transport emissions, reflected in policies at all administrative levels[1]. The transition to EVs hinges on a successful energy transition and vice versa: on average, EVs are connected but not charging 73% of the time[2], displaying potential for load-shifting[3] through vehicle-to-grid setups[4]. Indeed, as public charging infrastructure expands to match the EV adoption rates, EV charging behaviour emerges as a key social practice in shared urban contexts[5]. While EV charging behaviour has received attention with respect to typecasting of user profiles[6,7] or its cognitive dimensions[8,9], technical or modelling studies are typically system-focused[10] or use stylised assumptions[11,12]. Meanwhile, realistic behavioural dynamics remain understudied. In this paper, we address this gap through participatory inverse modelling.

Agent-based models (ABMs) are versatile modelling tools that aid the study of complex, dynamic systems such as the energy transition toward fossil-fuel abatement or mobility[13–15]. Key to an ABM's applicability – decision support, policy evaluation, or behavioural insights – is its validity[16], scrutiny of its mechanisms, behavioural grounding[17], and interpretability that allow for deriving qualitative insights[18]. While ABMs are gaining increasing attention, initiatives for their reuse are comparatively underdeveloped[19], and their behavioural grounding and realism are often limited[20]. For EV charging behaviour dynamics, ABMs would enable studying of emergent patterns through a set of agent-agent and agent-grid interactions, and the challenge lies in scrutinising interaction rules as well as interpreting the results.

Indeed, the potential of integrating psychology and behavioural factors into ABMs and transition research is large[17,20], and the lack of behavioural grounding or realism is discussed in many modelling articles. Particularly, habitual or repetitive behaviour is understudied in favour of one-time adoption behaviour[21]. Participatory modelling is a loose umbrella term for forms of stakeholder engagement in modelling endeavours that might rise to this challenge: Participatory modelling approaches form "second generation"[22] or counter approaches to the secluded, scientists-only analysis of systems and models[22]: "if such models are representations of aspects of reality, how can it be possible to build them without inputs from people who interact with the system in reality?"[23]

[1]Energy Transition Lab, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628BX Delft, Netherlands. [2]Department for Resilient Energy Systems, University of Bremen, Konrad-Zuse-Straße 8a (NEOS), 28359 Bremen, Germany. [3]Emile Chappin, Geeske Scholz and Gerdien de Vries contributed equally to this work. ✉email: l.m.schubotz@tudelft.nl

Decisions derived through participatory modelling are "generally"[24] of higher quality and less contested. Involving stakeholders early on refines problem scope and formulation ex-ante and increases (confidence in) the realism of the model "rules" as well as the relevance of the outputs[23]. Indeed, the benefits are clear and undisputed; however, replacing the paradigm of "employing mathematical optimization algorithms"[22] with participatory methods is "both refreshing and frustrating"[23]: First of all, participatory modelling processes may be intensive time-wise, and carries the risk of simply not matching the policy processes aimed to inform or even uncovering related but entirely new questions. Further, there may be power dynamics skewing the result, differing motivations for participating, contrary perspectives or conflicting bodies of knowledge, and the risk that essentially, "the model [built] to some extent always remains a black box", requiring the stakeholders who may invest a lot of time to trust the model and the researchers. Lastly, there is a need to present the model outputs tailored to stakeholders to translate and relate back to the social problem at hand[23]. Nevertheless, all literature reviewed finds participatory approaches worth the trouble; specifically, the higher quality of the resulting model, its legitimacy and acceptance as an output of a collective process and its relevance due to the stakeholders' social learning is remarked; the model can facilitate the latter as a "boundary object", i.e. intersections of social worlds that allow for translation and understanding between those[25].

Furthermore, the participatory multi-modelling approach may contain "sense-making" sessions focused on "translating the relevance and utility of the constructed boundary object ecology for their own organisation"[25], i.e., reflecting on and translating from the model (outcomes) into their own worlds and expertise. This interpretative ex-post step is remarkable as the participatory modelling literature reviewed focuses on building rather than interpreting models. Indeed, the model's intended use determines participant selection, implying their role and, therefore, the nature of the insights gained. To our knowledge, there exists no general review on participatory agent-based modelling that investigates what participants were involved in, when, or for what purpose. However, the participatory modelling literature reviewed above suggests a substantial under-representation of domain expertise used to strengthen the model's behavioural grounding or interpretation.

ABMs pose challenges due to the potentially large number of parameter settings to test and their inherent deep uncertainty. Additionally, structural variation, i.e., changing the assembly and construction of the model via other blocks of code, remains especially challenging to achieve[26] as conceptual assumptions are difficult to quantify and test methodologically[27]. For this, a new method has recently been proposed[28] to generate and evaluate a large number of structurally different models using inverse modelling: Conceptually, while the "forward problem" moves from given initial conditions to an unknown model outcome, the "backward" or "inverted problem" aims to infer the unknown initial conditions that best approximate a pre-defined model outcome. As such, instead of a search from cause to effect, inverse modelling is a search from effect to possible causes[29].

For agent-based models and social simulation specifically, inverse modelling has been undertaken[30], and more recently presented as Inverse Generative Social Science (iGSS)[29,31–34], which resembles symbolic regression[35] within ABM structures. These works introduced and demonstrated the broad applicability of the inverse method for ABM. The iGSS approach was developed further computationally[36] and with more attention to interpretation, or in another spin on inverse modelling related to but separate from iGSS, putting forward a generative method for structural model exploration by generating and evaluating permutations of function calls[28]. Despite their differences in approach, topic, and focal points, all these works subscribe to the generativist agenda[37]: growing a phenomenon means explaining it. Being in its nascent state, however, inverse modelling yet lacks a mature interpretative systematics to live up to its exploratory and explanatory potential for computational social science.

Based on the literature reviewed, we can identify three gaps worth illuminating and exploring in the context of EV charging. First, EV charging behaviour is an important, novel, and emergent part of the Dutch energy transition with understudied behavioural dynamics. Second, behavioural experts are underrepresented in participatory modelling efforts, and agent behaviour is rarely grounded in behavioural theory despite a large potential expected. Third, inverse modelling is a promising technique for exploring models and seeking explanations for observed phenomena by growing them; however, it yet lacks a framework for interpretative sense-making and theory building.

Therefore, our research aim is to combine participatory and inverse modelling techniques to study the dynamics of EV charging behaviour.

We introduce a participatory inverse modelling approach that braids qualitative expert co-design with quantitative inverse modelling, and we apply the approach to study the dynamics of Dutch EV charging behaviour.

Methodology

Our approach braids participatory and inverse modelling into a single methodology. Figure 1 shows the schematic workflow of our participatory inverse modelling approach. The approach is divided into ex-ante sense-making, inverse modelling, and ex-post sense-making. First, a case, an existing model, data and experts must be sought and selected to be mutually compatible. With the experts, the model, case, and data must be understood and interpreted, leading to the design of experiments to be conducted. This step forms the bridge over to inverse modelling as the designs originate from qualitative sense-making, but must necessarily be congruent with model implementation. The experiment determines the modelling task; model runs produce output to be analysed. The analysis is part of both inverse modelling and ex-post sense-making because the results must be presented in formats tailored to the participants to facilitate translation and relation to the case[23]. Closing discussion and interpretation synthesise insights to decide whether another iteration of the same process is necessary. Indeed, the methodology is inherently iterative due to the uncertain nature of participatory modelling, which may uncover related but entirely new questions during the process itself[23].

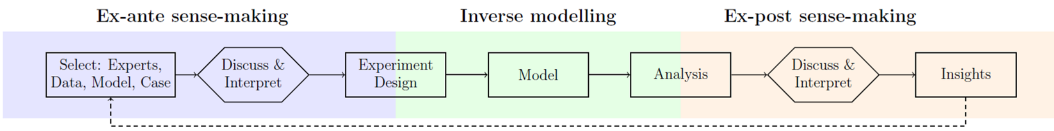

Fig. 1. Schematic visualisation of the participatory inverse modelling workflow. Hexagons represent those activities or steps including domain experts, rectangles those not including them.

For this research, we conduct two iterations of the workflow shown in Fig. 1: first, we generate and evaluate models with differing conceptual assumptions in structural experiments; second, we parameterise promising models. This split is to handle the computational demands of joint structure-parameter calibration. For the specific application of our method, our operationalisation, and model explanation, see Supplementary Methods S1. For a formalisation of agent state space representation and decision processes in the original model we refer to original model's the ODD protocol[38].

Formally speaking, let $\alpha \in \mathcal{S} \subset \mathbb{N}^S$ denote a multi-index for structural variation in $S \in \mathbb{N}$ instances as per the structure space $\mathcal{S}$ designed through domain expertise. Let now $\mathcal{M}^\alpha : \mathbb{R}^D \to \mathbb{R}^D$ the D-dimensional vector field that represents the base model $\mathcal{M}$ with structural configuration α mapping from the model's state space to itself, i.e. applying $\mathcal{M}^\alpha$ denotes one model step. For $n \in \mathbb{N}_0$, the $(n+1)$th state of the model $m^{(n+1)}$ is given iteratively through $m^{(n+1)} = \mathcal{M}^\alpha(\theta, \sigma^{(n)}, m^{(n)})$ for a fixed a parameter $\theta \in \Omega$ and initial seeding $\sigma := \sigma^{(0)}$ for the stochastic process. The inverse modelling process records the model output of $N \in \mathbb{N}$ model iterations through the projection $p : \mathbb{R}^D \to \mathbb{R}^d$ to the subspace of relevant model data, i.e. the sequence $(p(m^{(n)}))_{n=0}^N$ of model outputs is used to asses the loss of a single run denoted by $l(\alpha, \theta, \sigma) = \varphi(p((m^{(n)})_{n=0}^N))$ through an error function φ that compares the model output to the empirical data. The analysis estimates a loss $\hat{\mathcal{L}}(\alpha, \theta)$ and selects those tuples $(\alpha, \theta) \in \mathcal{S} \times \Omega$ that satisfy, first, $\min_{\alpha \in \mathcal{S}} \hat{\mathcal{L}}(\alpha, \hat{\theta})$ for a fixed, plausible $\hat{\theta} \in \Omega$ derived through domain expertise and, second, $\min_{\theta \in \Omega, \alpha \in \mathcal{A}} \hat{\mathcal{L}}(\alpha, \theta)$ for a subspace of well-performing structures $\mathcal{A} \subset \mathcal{S}$ that are pre-validated through domain expertise. These well-performing tuples (α, θ) are interpreted to explain how and why the model $\mathcal{M}^\alpha(\theta)$ grows the empirical data and to gain insights about the data.

Results & analysis

Following our workflow in Fig. 1, this section presents the results of our modelling experiments (Modelling results) and our joint sense-making conversations before (Ex-ante sense-making: design of the experiments) and after modelling (Ex-post Sense-Making: Interpretation).

Ex-ante sense-making: design of the experiments

From the interviews, it became clear that research in behavioural psychology typically captures static snapshots of behaviour, rarely formalised and often under the assumption of linearity, and it was considered interesting to explore alternative, non-linear, evolving dynamics[39–43]. The model provided a crucial context for studying the temporal dynamics of charging behaviour, and, thus, the goal of the sense-making process was to design and set up experiments to be implemented in the model. With our focus on behavioural dynamics and given this model, we selected Range Anxiety (RA) and Environmental Self-Identity (ESI) – whose weighted difference determines the agents' charging behaviour – as candidate variables for structural modification.

The domain experts based their expectations for the dynamics of incrementing and decrementing ESI and RA on social identity theory[44,45] (SIT) and system justification theory[46,47] (SJT), as both SIT and SJT are related to non-linear attitudinal inertia, with SIT focusing on group-bound threshold change and SJT on system-bound resistance. SIT was interpreted as for each individual to have a social identity and, relating to their in-group and making efforts to stay away from out-groups, resisted change under external pressure. Once, however, some threshold of pressure was passed, they would change their social identity, and one of the out-groups became their new in-group. SJT was described as a cognitive bias leading individuals to see the (position in the) system they were embedded in more favourably, thus displaying a resistance to change and external pressure to change. Similarly, these arguments would apply to decrementing the values, too; however, the "pull" factors that made the disposition (e.g., toward high ESI) favourable had no immediate counterpart as an ESI or RA value of zero meant ambivalence. In other words, based on SIT and SJT, a fade-out from extreme values was described as initially slow and then moderately accelerating. In summary, asymmetric nonlinearity of attitudinal change towards a value begins with inertia due to ambivalence and is followed by an accelerating shift after a psychological threshold is reached. Disengagement, in turn, occurs slowly at first and accelerates as attachment weakens.

Based on this analysis and the interviews, we designed alternative, non-linear curves for incrementing and decrementing the agent attributes range anxiety (RA) and environmental self-identity (ESI) that are summarised in Table 1. To update an agent's ESI or RA value along such an update curve φ, one would have to take the inverse of the function, apply the increment or decrement, and apply the function again, i.e., the new value is given by $y_{\text{new}} = \varphi(\varphi^{-1}(y_{\text{old}}) \pm \Delta x)$. Thus, the plot of φ in Fig. 2 acts as a reference chart; the update rule applies that chart to each agent's current state and moves the agent's attributes along the curve. Refer to the Supplementary Methods S1 for the modified model.

Following the logic of the theories, the domain experts expected the cubic function to accurately model the incrementation process for both RA and ESI. Furthermore, the modified sin function, scaled to match the respective boundary conditions, was expected to model the decrementation process for both RA and ESI. Lastly,

Label	Function	Responsiveness		Interpretation
		Middle	Extremes	
Linear	x	Moderate	Moderate	Nudge size equals effect size regardless of opinion
Cubic	x^3	Low	High	Hesitant, moderate
3rd root	$\mathrm{sgn}(x)\vert x\vert^{1/3}$	High	Low	Sticky and polarised
Sin	$\sin(\pi x/2)$	Moderate	Moderate-low	Moderate shifts, more sticky at extremes
Arcsin	$2\arcsin(x)/\pi$	Moderate	Moderate-high	Moderate shifts, less sticky at extremes

Table 1. Details on the update curves visualised in Fig. 2.

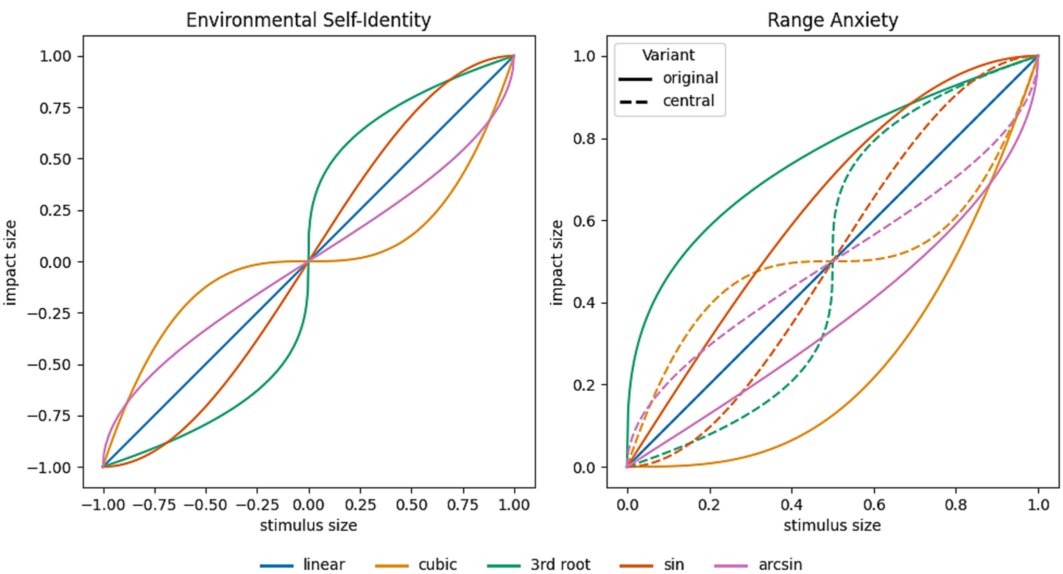

Fig. 2. Visualisation of the different update curves for both increment and decrement of Environmental Self-Identity (left) and Range Anxiety (right); for details, see Table 1. For RA, it was discussed whether ambivalence was to be set at 0, or whether it was to behave similar to ESI with a middle points at (0.5, 0.5), and the discussion about a "baseline range anxiety" sparked the design of "centralised" curves in Fig. 2.

other domain experts thought the third root function best captured the sensitivity of ambivalence: Agents with ESI or RA values close to zero would be more susceptible to nudges than those with extreme values in either direction. All these considerations are captured in the chosen alternative structures explored in the modelling.

The domain experts helped trim the parameter space size from originally 10^{24} down to 10^9 and select a single plausible set to use for the first iteration, the structural exploration; see Table 2 for details. Furthermore, all domain experts agreed that "losses loom larger", meaning that negative experiences have a stronger effect than positive experiences. This shared agreement has been operationalised in the parameterisation for the structural exploration with a higher RA increment and ESI decrement than the RA decrement and ESI increment, respectively.

Lastly, the domain experts found it interesting to investigate alternative permutations of the function call order in the models to gain insight into the underlying stage model. Indeed, prior work[28] found indications that structural order influences model results. See Supplementary Methods S1 for the original go-function to be varied.

Modelling results

This subsection states the results of the experiments designed before they are interpreted in subsection Ex-post Sense-Making: Interpretation. First, the results of the structural exploration are presented, where we narrow down on the top 45 model structures. For these well-performing structures, we explore which curves for ESI and RA increment and decrement specifically relate to a good fit. After remarks on the order of the structures, we conclude with parameter importances, partial dependencies of the errors between parameters, and structure-parameter interactions.

Structural exploration

Figure 3 shows three plots: the real charging data Fig. 3a in Den Haag during July 2024; the mean and standard deviation of the model output from the top 45 structures (Fig. 3b); and the mean and standard deviation of the model output from all tested structures (Fig. 3c). From visual inspection and juxtaposition, we observe

Variable name	Explanation	Range decision
f-esi-plus	Effect size of an increase in ESI which is updated during charging and decrements if the energy balance is negative, and increments else	{0.0001, 0.0005, 0.001, 0.005*, 0.0075, 0.01, 0.025, 0.05}
f-ra-plus	Effect size of an increase in RA which is updating during driving and decrements if the battery is zero and increments else	{0.005, 0.0075*, 0.01, 0.025, 0.05, 0.0625, 0.075, 0.0875, 0.1, 0.125}
f-esi-min	Effect size of a decrease in ESI	{0.0001, 0.0005, 0.001, 0.005, 0.0075*, 0.01, 0.025, 0.05}
f-ra-min	Effect size of a decrease in RA	{0.0001, 0.0005, 0.001, 0.005*, 0.0075, 0.01, 0.025, 0.05}
esi-mean	Average ESI over population (normal distribution)	{−0.25, −0.125, 0.0*, 0.125, 0.25}
ra-mean	Average RA over population (normal distribution)	{0.0, 0.05, 0.1, 0.15*, 0.25, 0.5}
esi-sd	Standard deviation of ESI over population	{0.05, 0.1, 0.15*, 0.2, 0.25, 0.3}
ra-sd	Standard deviation of RA over population	{0.05, 0.1, 0.15*, 0.2, 0.25, 0.3}
weighing-factor-esi	Weight of ESI in weighted difference with RA, which is calculated as a deciding score to determine the charging behaviour	{0.3, 0.35, 0.4, 0.45, 0.5*, 0.55, 0.6, 0.65, 0.7}
policy-intervention	Policy implemented	"information and feedback"*
ev-types	EV fleet constellation	{"fev only"*, "phev and fev"}
renewable-coverage	Coverage of renewable energy in the grid in percent	{0.25, 0.3, 0.35*, 0.4, 0.45}
solar-vs-wind	Ratio of solar to wind energy	{0.2, 0.3, 0.4, 0.5*, 0.6, 0.7, 0.8}
unlimited-charging?	Toggle for every destination to have an available charging point	{True*, False}
social-charging?	Toggle for agents to move their vehicle once charged	{True, False*}
central-control?	Toggle for a central control system	False*

Table 2. Parametrisation ranges decided on for the parameterisation based on interviews or estimated as properties of the region. The superscript * indicates that this value has been used in the structural exploration. ^{a}It turned out that the "unlimited-charging?" parameter was actually limited to 78, the number of workplace chargers.

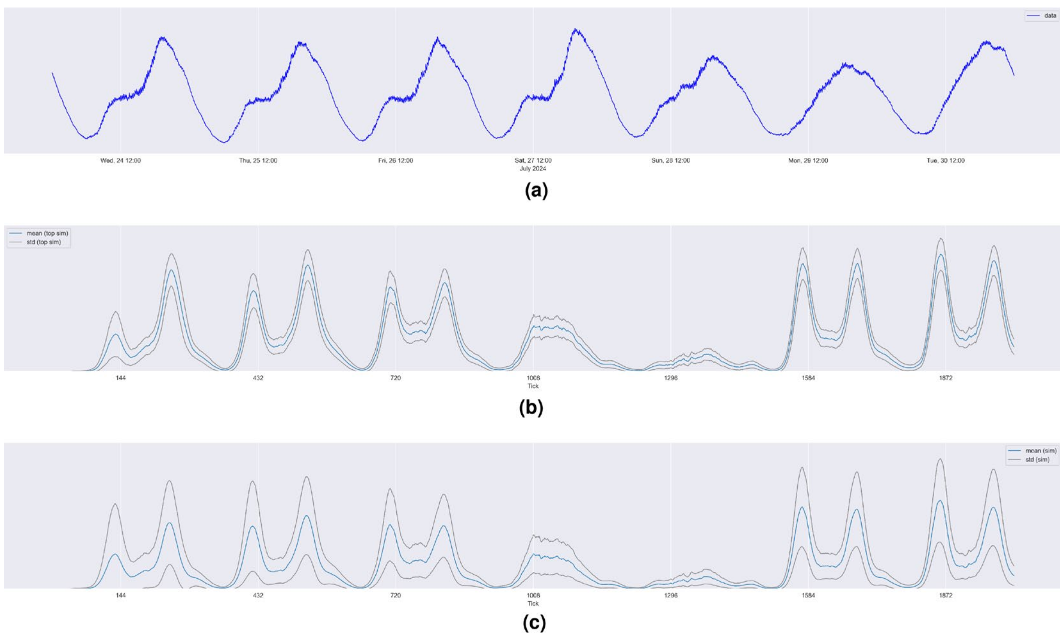

Fig. 3. Comparison of real data (**a**) with modelled charging demand from all structures explored (**c**) and top-performing structures (**b**).

the following differences resulting from seed and structural variation: The real charging demand shows a staircase shape with a saddle point just past noon and a peak in demand in the evening. The charging demand from Wednesday to Friday is regular and continued on Saturday. In contrast, from Sunday to Tuesday, the real charging demand is lower and steadily increases until the evening peak, rather than the previously observed staircase. By contrast, the simulation data in Fig. 3b and Fig. 3c differ from the real data: All weekdays are regular and congruent with one another. Instead of a staircase shape, the simulation data shows peaks past noon and in the evening, separated by a minimum. Weekend charging demand is lower without identifiable peaks. The weekday pattern resumes after the weekend. The model output of the top 45 structures differs from that of all the structures insofar as the minima between noon and evening peaks are more elevated in Fig. 3b than Fig.

3c, matching the saddle point shape in the real data more closely. Second, the top simulated charging demand shows higher evening peaks and a much tighter standard deviation band. Particularly on Saturday, the model output shows a higher charging demand, whereas Sunday demand Fig. 3b resembles Fig. 3c. In summary, while the model accurately matches nighttime demand, it fails to match charging demand, particularly during lunch breaks, exhibiting two distinct peaks rather than a saddle point near noon, with a peak toward the evening, even for the top 45 model structures. Furthermore, weekends and weekdays differ greatly as weekend charging demand is barely existent in the model and the regular weekday pattern continues afterwards.

The comparison in Fig. 3 highlights the potential and limits of our experiment design for structural exploration that is informative despite being conducted with one single parameter set: Structure matters, and we can select and identify structures with higher quality of fit – robustness through tighter standard deviation bands, reduced error through elevated midday minima, and improved fit on Saturdays – purely through alternative behavioural update rules. However, for this parameterisation, none of the explored behavioural designs fully overcame deviations between weekends and weekdays, post-noon minima, or inter-weekday irregularities. These are inherent to all runs and suggest being schedule or profile-dependent[48,49], and therefore model architectural or conceptual in nature beyond our experiments; we will return to this point in the discussion.

Increment and decrement of RA and ESI

For the individual results of the increment and decrement curve of RA and ESI, consult Fig. 4, which contains the increment and decrement curves for RA and ESI that were found in the top 45 model structures. For the RA increment, the third root was not observed, neither normal nor centralised, indicating that RA grows steadily continuously with no threshold-based jumps. For the RA decrement, all curves except the third-root occur in the top 45 structures, indicating a linear or sublinear process in which extreme RA does not persist for long, but a slight fear always remains. The ESI increment showed that all curves yielded similar results in both metrics, so no pattern could be identified. The ESI decrement, in turn, lacked the arcsin curve, and the errors for the cubic decrement were high, suggesting that the reduction in ESI is moderate and that the positive extreme is rather sticky, with no plateau of ambivalence.

Lastly, from a cross-comparison, there was a clear best combination for both mean-squared error (MSE) and dynamic time warping (DTW): increment RA cubic, decrement ESI third-root, and increment ESI "superlinear" (i.e. linear, sin, and third-root), and decrement RA inconclusive.

Structural order

Regarding the order of function call execution, we examined the order of the top 45 structures. However, hardly any emerging dynamics could be traced through interaction. Preliminary testing suggested fixing `setup-for-step` and `electricity-demand-and-supply` at the beginning and `output-indicators` and `tick` at the end for the model to compile. Analysing the best structures, we found that `adopters-travel` was always called before `set-destination`. Investigating the code, we found that `adopters-travel` provided the necessary day profile at midnight for `set-destination` to use for the next day. Experiments confirmed it to be irrelevant where `adopters-travel` was placed as long as it was before `set-destination`. However, no further order-related patterns could be identified. This suggests that the *order* of function calls is irrelevant, but rather the composition, i.e., *which* functions or update curves a structure contains.

Parameterisation

Concerning parameterisation, Table 3 lists the parameters sorted by influence based on our evaluated sample. For both metrics, the parameters `ev-types`, `solar-vs-wind`, and `esi-mean` (and potentially `renewable-coverage`) appear to exert a strong influence on the quality of fit. Overall, we see that contextual parameters show a stronger influence than behavioural parameters, except for `esi-mean` and `social-charging?`. These results will be interpreted in Parameterisation.

Dependence and interactions

We control for parameter interactions that could distort the error values; see Supplementary Methods S1. The partial dependence plots show merely sporadic interaction effects accounting for less than 1.38% of the output range in the strongest case, which can be considered weak to weakly moderate. This interaction, however, between `solar-vs-wind` and `renewable-coverage` is detected in all 45 models; the model code indeed shows a linear, multiplicative scaling of the two variables. This textbook case of statistical interaction serves as internal validation of our method: the SHAP analysis reproduces the confirmed findings, and the metrics MSE and DTW are sensitive enough to detect these interactions. Other interactions with lesser effect sizes were investigated, but no link could be established, increasing confidence in the best values (Table 3). Lastly, for this well-performing sample of structures, no notable structure-parametric interactions were detected, meaning the tested parameters had very similar effects regardless of the structure, and that for these well-performing structures, structure and parameterisation were quasi-separable.

Ex-post sense-making: interpretation

This subsection contains the discussions held and explanations provided to explain the results stated earlier.

Structural exploration

We could explain the mismatch between real and simulation data in Fig. 3 through several aspects. First, regarding the two-peak shape of the model data, we found that agents would always plug in their car when the state of charge was less than 100%, whereas people would not necessarily recharge immediately or until

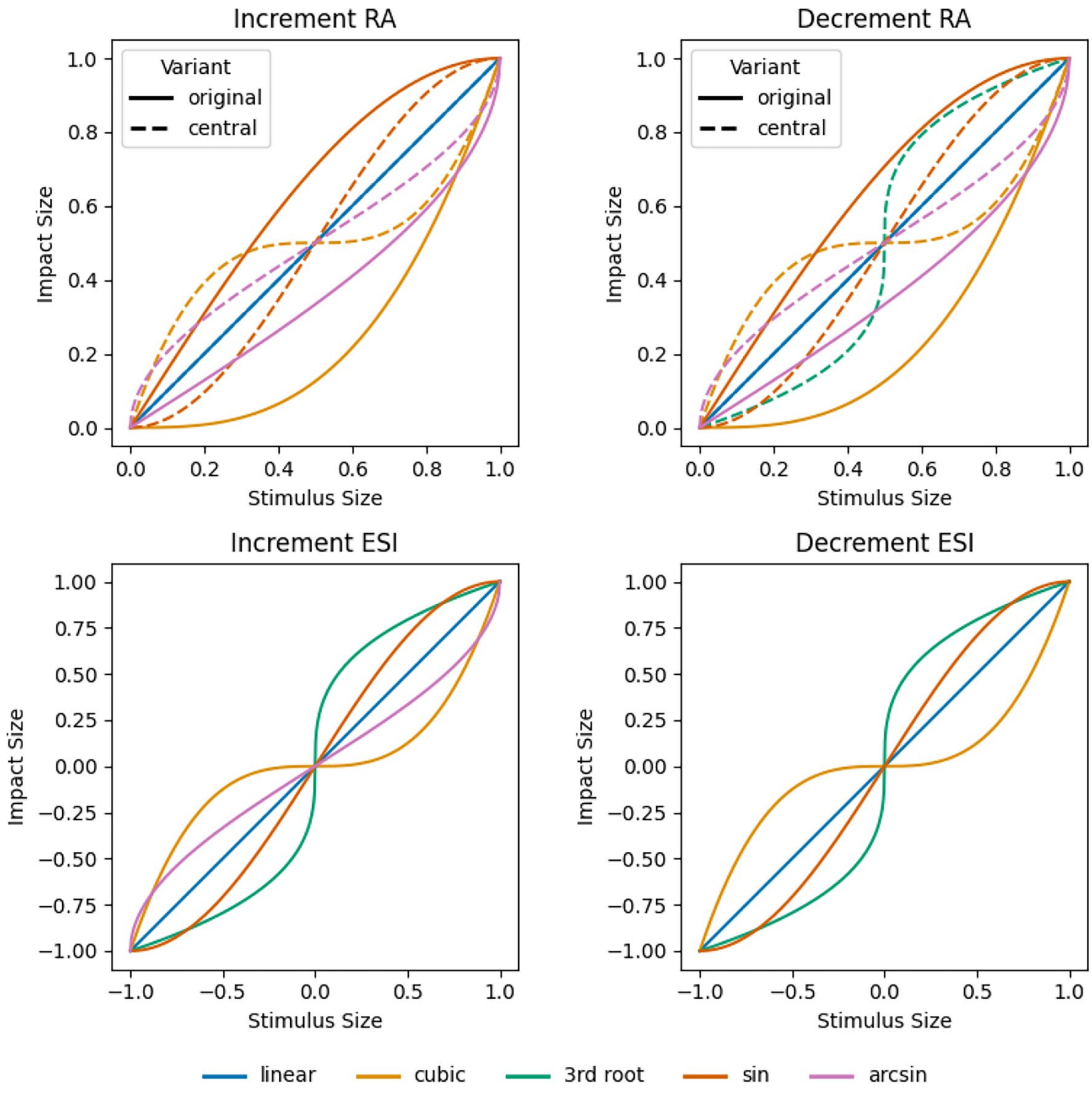

Fig. 4. Results for observed increments and decrements for RA and ESI in the top 45 structures. The best combination found consisted of cubic RA increment, third-root ESI decrement, and "superlinear" (i.e. linear, sin, and third-root) ESI increment, with RA decrement inconclusive.

full[7,50]. Short morning commutes resulted in low charging demand until past noon, thereby explaining the bimodal shape of the model's aggregated charging curve. Second, the striking difference between modelled and observed weekend demand was due to profiling and scheduling, as all agent activity besides work, shopping, and being at home was categorised into one of these three categories[12]. Specifically, they counted leisure activities as empty days for the agent. Meanwhile, the domain experts argued that weekends were less regular than the workday schedule, particularly Saturdays, which typically involved multiple trips of varying distances. Lastly, the discrepancy between Wednesday to Friday compared to Monday and Tuesday was perceived as bewildering by all domain experts and could only be hypothesised to be due to work week irregularities – the Netherlands being among the highest-ranking countries in Europe with respect to part-time work or work from home – or the fact that the data for the model was aggregated from all of the Netherlands from before COVID, but the charging data utilised to match was from Den Haag only and from 2024.

Increment and decrement of RA and ESI

We presented the results, summarised in Fig. 4, to the domain experts.

The RA increment results showed no third root, neither normal nor central, which we interpreted as RA growing continuously but not abruptly after some threshold. Based on this, the domain experts thought RA was conceptualised as a trait in the model but wondered why RA, as a trait, would change over time scales of less than a month. Whether RA was a trait or a state, and whether RA was related to the trait of anxiety, was unknown and deemed worth further research. The RA decrement results showed no third root, which we interpreted as indicating that RA reduction is a linear or sublinear process. In other words, high RA does not reduce slowly,

Parameter	Avg. rank		Avg. |SHAP|		Best range
	MSE	DTW	MSE	DTW	
ev-types	1.44**	1.0**	0.081	0.125	"fev-only"
solar-vs-wind	1.6**	2.1**	0.065	0.068	{70, 80}
esi-mean	4.4**	4.2**	0.032	0.04	{−0.25, −0.125}
renewable-coverage	4.4**	5.3*	0.028	0.029	{0.25, 0.3}
f-ra-min	6.6*	6.2*	0.021	0.029	{0.0075, 0.01, 0.025, 0.05}
unlimited-charging?	6.5**	7.1*	0.019	0.018	False
weighing-factor-esi	7.6	6.9*	0.016	0.021	{0.3, 0.35, 0.4}
f-esi-plus	8.9*	8.9*	0.016	0.02	{0.0001, 0.0005, 0.001}
f-ra-plus	8.9	8.6*	0.012	0.015	{0.05, 0.075}
ra-mean	9.2	8.8	0.011	0.014	{0.25, 0.5}
f-esi-min	10.0	10.4	0.009	0.01	{0.025, 0.05, 0.075}
esi-sd	10.4	10.4	0.008	0.009	{0.3}
ra-sd	12.6	12.4	0.003	0.004	inconclusive
social-charging?	12.5	12.7	0.003	0.004	False

Table 3. Comparison of features by MSE and DTW metrics, sorted by average rank of importance over all models. In terms of notation, the superscript ** indicates importance for the majority of models in MSE/DTW (more than 25 models); * indicates importance for some but less than the majority (more than 10 models); absence of a letter indicates no importance for that metric. [a]In machine learning, "importances" refer to how much each "feature" (parameter) contributes to the prediction accuracy of the model. Mean absolute SHAP values lack a standalone interpretation but can be understood relative to one another.

and low levels persist. If RA were an emotional state, these results made sense to the domain experts, who expected a third root if RA were a personal trait. Originally, the centrality curves were designed for RA as a state, but no pattern could be identified. As a general note on RA, the domain experts argued that the existence of RA was schedule, memory, and context-dependent: 30% of battery could be enough to get to work and home again, where a charger was nearby, but not enough for irregular Saturday tasks with uncertainty about charging availability. Once routine around new tasks had settled in, RA related to the specific schedule and context would decrease, but different tasks with uncertainty would cause RA to spike again – this was at odds with the implementation in the model. As a final interpretation of RA, the domain experts interpreted the model design – agents looking to charge whenever their battery was not full – as an assumption for a preference for certainty. Drivers would avoid the risk of running low by restoring maximum range whenever the opportunity arises, rather than optimising charging timing. Conversely, the model distinguishes this from situations of necessity: when the battery is empty, and charging is unavoidable, behaviour is driven by constraint rather than anxiety, and therefore does not meaningfully affect preferences or identity as discussed above. In behavioural terms, RA can be seen as a preference for charging security, i.e. a reliable and predictable access to charging. This need for certainty can also manifest as charging earlier or more fully than strictly necessary.

The ESI increment results showed all curves present with similar results for both metrics. The domain experts were surprised, but thought that "more complexity is always good in terms of identity" and that the result helped to argue for other underlying influences beyond our experiments or the model itself: effect sizes in psychological research were generally low compared to STEM fields[51] and variables of low influence better replaced or modified, e.g., in this case charging spots availability from a true-false switch to a variational parameter range. Indeed, ESI was a broad psychological construct that, in the model, served as a preference for renewable energy, operationalised in a postulated feedback loop between preference, the availability of excess energy, and behaviour. Given the intention-behaviour gap, however, a mere preference was deemed insufficient for such a hypothesised feedback loop. The ESI decrement results showed no arcsin and a poor fit for the cubic decrement, which was interpreted as strong identities being resistant, and that the region of ambivalence is not a plateau. Seeing ESI as an individual's preference, the result was plausible as ambivalence was a state that could easily be swayed one way or another; if ESI was conceptualised as a norm, however, they would expect the opposite result as strong norms could easily be weakened, but would still be upheld weakly for a long time. They noted further that the result was based on the assumption of no behaviour-intention gap, since in the model, ESI as a preference had an immediate impact on the agents' charging behaviour. According to the well-known COM-B model[52], however, it was capabilities, opportunity, and, most weakly, motivation (via intention) that influenced behaviour. Going from preference via motivation to behaviour was seen as a leap. As a final interpretation for ESI, the domain experts discussed, similar to RA, whether ESI was more of an identity (as per the name) and thereby a temporarily fixed trait or a preference (for renewable energy as operationalised in the model) and thereby a state. Again, they expected more sublinear curves (cubic, arcsin) for a trait and superlinear (sin, third root) for a state. However, the extent to which identity was related to ESI and, in turn, to a preference for renewable energy remained an open question.

The best combination suggested the following dynamics to be at play: cubic RA increment was the fastest possible growth given a sufficiently high base level; third-root ESI decrement was stable for sufficiently high identities but also quickest reduction for even some positive values; "superlinear" (i.e. linear, sin, and third-root) ESI increment, in turn, reflected fastest possible growth but stable negative identity. The inconclusive RA decrement could not be interpreted further.

Structure order

The domain experts were not puzzled that it was not the structural order that seemed to matter, but rather its composition. Stage models, they explained, progressed by passing thresholds sequentially, which, when shuffled, would indeed seldom remain meaningful. By contrast, continuously interacting designs – such as the aforementioned COM-B or similarly well-known UTAUT model[53] of technology acceptance and usage behaviour – have increasingly gained attention and recognition. In this view, the ABM's architecture was somewhat hybrid: while the functions were called sequentially, they interacted with each other, as charging decisions depended on RA and ESI values updated by other functions. However, thresholds were also part of the model, inherent to the functions themselves; for example, the charge mode was determined by whether the weighted difference between ESI and RA exceeded a threshold; see the original ODD protocol[38].

Parameterisation

First of all, the universal strength of the influence of `ev-types` is noteworthy as a heterogeneous fleet had not been considered earlier[12]. Based on this and our discussions with more technically versed domain experts, we decided to use a fully electric fleet for the structural exploration. Explaining this universal strong influence, however, we found that the setting "fev only" results in higher charging-demand peaks than a hybrid fleet does. Our prescaling after the structural exploration, in turn, was tuned to the setting "fev only", and it is therefore unsurprising that the same prescaling with another fleet composition would yield a worse fit. For a different fleet, another prescaling might give different results.

`solar-vs-wind` had priorly[12] been found insignificant for the results and excluded for `renewable-coverage`. We detected a linear interaction between the two parameters (see Supplementary Methods S1) and found both to be important for explaining the goodness-of-fit. Indeed, Table 3 not only displays both parameters as crucial but shows that `solar-vs-wind` around 70 − 80% and `renewable-coverage` around 25 − 30% give the best fit to the real data. The model uses weather data from the Belgian coast in 2016, and January is characterised by few hours of sunlight and strong winds. The values in Table 3 of low renewable coverage and solar dominating over wind yield practically no excess energy, leading the agent's ESI to decrease when charging, in turn, leading the agent to most likely end up in charge mode 1, i.e., to always charge until the battery is fully charged regardless of the grid. Even though the settings of low renewable coverage and high solar do not correspond to reality, these settings yield the best fit under the assumption that charging relies on RA, leading us to conclude and conjecture that charging behaviour in Den Haag is mainly driven by RA or, in turn, by a lack of consideration for the availability of excess energy. This model-based finding may manifest by various circumstances in public charging settings: users may either disregard it to fulfil their need charging security or possess no means to access this kind of information, possibly through grid-sensitive vehicle-to-grid charging modes to automate this process[4]. Regardless, this conjecture is further strengthened by `ev-types` being fully-electric vehicles only, effectively increasing the charging demand, as well as `esi-mean` being negative and `weidhing-factor-esi` being low, favouring RA in the decision about the charge mode.

Regarding the behavioural parameters, note how `f-ra-min` and `f-ra-plus` are of roughly the same magnitude, with `f-ra-min` being slightly lower, and, further, how `f-esi-min` is between 25–750 times as high as its incrementing counterpart, `f-esi-plus`, which took the lowest possible values of the available parameter grid. The domain experts related this to prospect theory[54] in behavioural economics; the classical result is that, in a 50:50 chance game, the potential gain must be about twice the potential loss to accept the odds, showing that losses loom larger than gains. However, while the value of prospect theory was seen in its quantifiability, generalisations beyond finance-related behaviour had to be taken with appropriate caution. Therefore, since the behavioural grounding of the designs was also possible through SIT and SJT, prospect theory had been discarded earlier. In turn, the parameterisation results indicated that RA and ESI were implemented as states rather than traits. However, specifically for ESI, an identity, they would have expected a more stable identity. While preferences were also subject to change, they were reported to be stable when the decision was important and tied to personal values. Therefore, based on the model and parameterisation results, they interpreted EV charging and use as economic decisions rather than as enacted environmental identity, and argued that climate intentions did not lead to climate actions[55].

Regarding structure-parameter interactions, while no significantly distinct parameterisation could be identified for the top 45 models, the results of the structural analysis support the interpretation of the parameterisation: Notably, our hypothesis is in line with the best found structural combination: cubic RA increment was the fastest possible growth given a sufficiently high base level that is indeed given with best-fit values for `ra-mean` of 0.25 or 0.5. The third-root ESI decrement was interpreted as stable for sufficiently high identities, but also the quickest reduction, even for some positive values. A best-fit value of −0.125 or −0.25 for `esi-mean` with `esi-sd` of 0.3, despite its low rank and importance, points towards steep ESI descent. The inconclusive RA decrement could not be interpreted further; the "superlinear" (i.e. linear, sin, and third-root) ESI increment, however, points toward linear to steep increases for positive ESI values, with merely gradual increases for negative values in turn. Particularly, the fact that the decrement, `f-esi-min`, is between 25–750 times as high as its incrementing counterpart, `f-esi-plus`, supports that the best-fit structure-parameter combination drives charging behaviour mostly through high RA and negative ESI.

Discussion

This research aimed to combine participatory and inverse modelling techniques to study the dynamics of EV charging behaviour. We operationalised this by using an existing model and had inverse modelling infer structural and parametric conditions to match EV charging demand data from Den Haag. We involved generally under-represented behavioural experts in EV charging (behaviour), energy transition, and social simulation to design and interpret behaviourally grounded alternative, non-linear dynamics. We then tested and analysed these dynamics through inverse modelling. So what? Combining inverse and participatory techniques in depth is a novel approach. It allows creating, testing and scrutinising realistic behavioural dynamics that were previously unattainable, and sharpens questions for future behavioural psychological research. Furthermore, it enhances explainability and increases trust and transparency in the model's capabilities, limitations, and potential areas for future improvement.

Discussion on results and behavioural modelling

Firstly, our experiment designs in Fig. 2 and Table 1 show that behavioural grounding of modelling choices can be achieved. Secondly, Fig. 3 displays the importance of structural variation that has previously been left unaddressed in ABM endeavours. Thirdly, we have tested for order, partial dependencies, and parameterisation by separating structure and parameters into a two-step workflow, and our analysis shows that these are meaningful endeavours: modelling choices, structures, parameters, and interaction effects can systematically be tested, analysed, and subsequently evaluated with behavioural expertise so that we can give meaning to these results. We were able to explain what parameters yielded a good fit and that the model matched the data best when the agents' charging behaviour was driven by the RA variable. This led us to conjecture that charging behaviour in Den Haag was mostly driven by range anxiety or, due to various potential reasons, a lack of consideration for the amount of excess energy in the grid. Although the model is not a hyperrealistic version tailored to Den Haag, with agent-internal processes and parameterisation aspects differing from reality, inverse modelling enabled us to select those well-performing designs and interpret them in a participatory process.

Furthermore, regarding our case and the behavioural dynamics, we have uncovered a range of considerations and questions that drive model development and behavioural research on this topic further. Firstly, in the model, the variables RA and ESI were operationalised inconsistently, rather as states than traits, and it is unknown which is true. RA was argued to be rather schedule, memory, and context dependent, with an under-researched link to anxiety traits. ESI, in turn, was argued to be a broad psychological construct that functioned not as a stable identity but rather as a preference for renewable energy. With the presumed functional relationship $\omega\mathrm{ESI} - (1-\omega)\mathrm{RA}$, our result that the agent variable RA is decisive of the charging behaviour can be turned around into a lack of ESI or, seen as a preference, a disregard or lack of consideration for the availability of excess energy in the grid. While this may be due to lack of information, choice, or care, these variables, ESI and RA, are fundamentally different in a psychological perspective, and $\omega\mathrm{ESI} - (1-\omega)\mathrm{RA}$ effectively weighs and subtracts semi-stable preferences and emotional state reaction about expected capabilities. Model implementation, therefore, needs to be specific: If a systemic psychological perspective on EV charging is taken, complex constructs like ESI must not be operationalised merely as a preference that, e.g., skips the behaviour-intention gap. This requires stronger, more interdisciplinary collaboration to translate psychological constructs into sound formalisations[41,56] and embrace non-linearity[43].

Our findings emerge from a participatory structure-parametric inverse modelling process that calibrates the model to empirical charging demand data. The results are first and foremost model-based findings and require critical reflection. Our result that agent charging behaviour is driven by range anxiety and with a lack of consideration for grid conditions is inherently an artefact of the model's mechanisms and assumptions, as we derived: low renewable coverage and high solar shares result in low excess energy, letting overall ESI shrink and making high range anxiety values dominant in the weighted difference, which decides charging behaviour. This constitutes candidacy for a plausible explanation for the observed real-world behaviour, but an explanation that is congruent with empirical findings: studies suggest that charging behaviour is driven by range anxiety[4], a multitude of interacting factors other than range anxiety[48], or range anxiety in relation to infrastructure availability[49]. Second, the availability of charging infrastructure has been shown to play a role in reducing range anxiety[49,57], showing that these two variables are intertwined. From our analysis, we found no interactions for the unlimited-charging parameter, and our model setup did not allow for granular scaling of infrastructure conditions. While simulations specifically enable the controlled study of the effects of infrastructural conditions beyond Table 2 (and footnote a), this would constitute a different study and the distinction whether users charge due to uncertain access or range anxiety is left for future modelling and behavioural research.

Our result that EV charging is largely driven by range anxiety or a lack of consideration for (or awareness of) the availability of excess energy in the grid is valuable to consider for policy. First, the aforementioned reciprocal relationship between range anxiety and charging infrastructure warrants joint consideration. In practical terms, policies must prioritise improving technical aspects such as charger reliability and availability, but rather than relying primarily on incentives to align charging behaviour with renewable energy generation, relevant information must be made available to users such as transparency of access, real-time information, and the option for grid-aware charging. In a similar vein, social norms can be leveraged to allow users to achieve charging security (i.e., reliable and predictable access to charging) by charging enough rather than charging to full[58].

Methodology employed

Our workflow, visualised in Fig. 1, proved suitable for our research aim to combine participatory and inverse modelling techniques to study the dynamics of EV charging behaviour. While we performed two iterations of the workflow, once for structural and once for parametric exploration, it would be entirely possible to perform another structural exploration using the newly found, best-fitting parameter set and so on. Indeed, our method is inherently iterative, allowing for testing and scrutinising behavioural theory and formalisation in constant exchange between modelling and domain expertise. Our analysis enabled us to identify not only the order but also the composition of an influential structure, and to select a small, well-performing subspace of the structure space where structure-parametric effects are negligible. However, it remains entirely plausible that strong structural-parametric and ordering coupling exists for other models, as prior findings suggest[28]. Therefore, and in line with the pointers on model improvement, our method is not only strong but also exhibits limitations: the workflow does not function as a black-box answer machine, but rather requires interactive and iterative human validation and expert reasoning and design[28]. Through this, scrutinised behavioural theory, improved model transparency, and a generative transfer back into social science can be achieved.

The participatory process

In line with prior work[23], the participatory process was time-intensive and required trust in both the model and the researchers. The domain experts helped refine the problem and research scope, as well as the research's relevance to their respective disciplines. While not per se conflicting, somewhat contrary information within that same body of knowledge was encountered, as, e.g., entirely opposite curves were deemed possible and could be behaviourally grounded. Lastly, particularly the early sense-making talks in which the project was outlined, as well as the ex-post sense-making sessions, required us to present the model outputs tailored to domain experts' needs. Indeed, our sense-making talks repeatedly focused on understanding each other: differences in notions were well known[59], we also think the specificity and formalisation of the model is at odds with the flexibility and broad applicability of the model: "It's very hard to think about it this way!" said one domain expert when pressed for their thoughts on how range anxiety would grow. In a similar vein, interpreting float values for parameters proved challenging. Despite the difficulties both researchers and domain experts experienced, they were excited about the updates on this research's progress and thought the contributions extended far into their respective domains. Indeed, the model acted as a boundary object for mutual teaching and learning[25]. Regarding future research, recent developments in participatory model-based research might prove complementary synergies to our approach: For similar application contexts[60], Large Language Models (LLMs) have been leveraged as synthetic stakeholders for planning processes. Our own use of domain expertise to evaluate and interpret model structures and the runs of an inverse workflow was important for assessing plausibility and critical for generative purposes such as hypothesis generation or theory building. However, given our time-intensive process and previous findings from our literature review – that domain expertise is seldom represented in participatory modelling efforts – the automatic, qualitative, and LLM-supported evaluation of model runs, supported by human domain expert input or control, may well enable a closer interaction of domain expertise with the modelling process and mid-modelling evaluation in future research. Tools like ASReview[61] might prove helpful to effectively judge the quality of these simulated expert assessments and ease the demands on domain experts over multiple iterations of modelling and consultation.

Inverse modelling

Inverse modelling proved crucial in this research, enabling targeted model exploration to match given empirical data. For this, random mutations and random parameterisation were used as a simplifying choice to achieve a broad enough sweep. While a fair amount of repetitions has been done based on a heuristic that balances computational load with coverage of the search space, distribution-agnostic sample sizes and search strategies that produce dense enough samples, both structural and parametric, were outside the scope of this study. For inverse modelling, however, confidence in prior results is vital for the success of subsequent guesses, given the potentially high cost of evaluating an ABM multiple times. This makes uncertainty quantification and balancing computational load a key challenge, and it is unclear whether intricate search and optimisation methods can solve it. Specifically, the metricisation of the "rule space"[29], which is comparable to our structure space, is challenging. While the distance between structures can be quantified[28], this shows large discontinuities in the structure-fitness landscape.

Our joint structure-parametric interpretive approach is an advancement of inverse modelling as a methodology and toolkit for generative ABM-based studies. Yet, other methods, such as Inverse Reinforcement Learning (IRL), exist for similar problems, e.g., sequential rule-based decision-making for travel itineraries[62], and it is worth relating and distinguishing these two methods. Both iGSS and IRL share the assumption that observed human behaviour is near or somewhat optimal with respect to some function of model variables, and so iGSS studies[29,31–34,36,63] focus on building fitness functions for macro-level metrics through genetic trees akin to symbolic regression[35]. Meanwhile, our approach[28] differs from iGSS in that we assume the observed macro-patterns emerge from a latent structure-parameter combination, where structures and parameters are not necessarily optimal but plausible given domain expertise. We then find optimal parameters for a multitude of plausible structures, validated by domain expertise, and leverage them to interpret these parameters. This is a "plausibility first" paradigm rather than searching for plausibility among optimised solutions. Second and regardless, both approaches share the inverse automated workflow[62], Algorithm 1 – building generative surrogate models, inferring a latent structure, and evaluating and interpreting the data – as our formalisation of Fig. 1 shows: IRL distils a policy π (similar to our structure-parameter combination (α, θ)) and analyses the reward R (similar to our loss $\hat{\mathcal{L}}(\alpha, \theta)$) through a multi-nominal logit model, and our approach evaluates structure-parameter combination through fitness and if well-performing, a final, qualitative interpretation via domain expertise as part of the modelling loop. Third, differences can be found in the model and data used: ABMs are similarly rule-based as utility-based decision models, but exhibit emergent properties with sensitive dependencies on initial conditions and seeding, making them computationally challenging to handle and interpret. In turn, the data (35,714 users over one day in 15-minute intervals) allow IRL to model individual decision-making and, through sequences of activity, validate at the microlevel and combat overfitting, while our approach identifies candidates for macro-level explanations from potentially sparse or aggregated, but longitudinal, data. Lastly, this discussed proximity in approach and capabilities offers interesting ways forward for future research: The connection of inverse modelling in social simulation to machine learning is heavily under-explored with few exceptions[31,36,63,64], and explainable AI techniques might aid to unlock this "generative"[29,37] potential. Using IRL for AMBs could solve many computational problems that inverse problems in social simulations face, such as improved scaling for search in high-dimensional parameter spaces. However, this requires differentiable spaces, which is challenging, firstly, as our structural search (the rule space is challenging to metricise[29]) represents different model assumptions or alternative system design categories, and secondly, for arbitrarily precise parameter values, which may be challenging to interpret. Therefore, hybrid approaches for separated structure-parameter search for already plausible structures would certainly be worth pursuing.

Participatory inverse modelling
The method we proposed combined participatory and inverse modelling techniques into a single methodology. Arguably, our participatory sessions functioned as a qualitative inverse process, in which we iteratively developed experimental designs and evaluated their fit to the research aim, model, and behavioural theory and grounding. In turn, computer-based inverse modelling was not separate but connected to the sense-making process. Therefore, our method braided qualitative participatory methods with quantitative inverse modelling techniques, and future research should leverage the discussed methodological and LLM-driven developments[60,62] in participatory and inverse modelling research.

Further limitations

Naturally, there is a risk of overfitting, since only one dataset was used to compare the simulation data. Therefore, no generalisability of our findings is given. Regardless of that, our method demonstrates that models and data can be made sense of with and through domain expertise.

Our analytical method was simple but sufficient. This approach has been favoured over Sobol, Morris, or Delta analyses using SALib[65,66] for several reasons: first, our aim is to explain goodness-of-fit rather than to estimate variance-based indices, as Sobol indices would permit. Second, the sample was not Saltelli-distributed but rather random, as a simplifying choice and to avoid biasing the parameterisation results due to structural influences. Thirdly, alternative methods were computationally too expensive, unlike SHAP values. Still, in relation to our complexity elaborations above, future or similar work should consider Morris or even Sobol analysis, despite their additional computational costs.

Lastly, we utilised two distinct metrics to evaluate the goodness of fit of a single run, with no global constraint on structural exploration and a Sakoe-Chiba band for parameterisation results. Using two metrics is novel and added both robustness and nuance to the analysis. We found that a metric is necessarily and always a lens on the data that forgives some aspects but punishes others: In our case, MSE accounted for exact point-by-point distances that asserted exactness, while DTW first aligned the time series like a smoothener and then calculated MSE of the aligned path. We found that DTW is more lenient in detecting minor importances. However, issues like dependence on prescaling persist, and the need for nuance across multiple metrics forces the researcher to make less clear-cut decisions in the analysis.

Conclusion

Our research aim was to combine participatory and inverse modelling techniques to study the dynamics of EV charging behaviour. To this end, we designed alternative, behaviourally grounded, non-linear dynamics, which we implemented and tested within an existing model. Drawing on relevant domain expertise, we illuminated the gap between inverse modelling and interpretative analysis, leveraged underrepresented behavioural expertise in participatory modelling contexts, and studied large-scale behavioural dynamics of EV charging rather than individual snapshots. The outcome is a general epistemic workflow for behavioural inference in agent-based models under structural-parametric uncertainty, enabled by combining participatory and inverse modelling.

We have shown the applicability of our method through a case study on EV charging behaviour in Den Haag. By matching simulation output to public EV charging demand over a week, our analysis identified a set of structural alternatives and parameters that match our Den Haag data well, with some caveats remaining. We discussed our findings with the domain experts and gave meaning to the results through discussion. We have gained a deeper understanding and have mutually taught each other using the model as a boundary object: we understand why these model structures are parameters that yield good results and have opened the black box of the model, both quantitatively and qualitatively. Our finding that modelled EV charging is largely driven by range anxiety or a lack of consideration for the availability of excess energy in the grid is valuable to consider for policy, and by discussing whether range anxiety and environmental self-identity are states or traits and therefore subject to what form of change, we have generated hypotheses forward for behavioural research to investigate further and considerations for model building and improvement.

While external domain expertise verified and validated our hypotheses and method enough to warrant serious consideration, the generalisability of our behavioural results is limited by case-specificity. Our workflow, however, stands as a methodological contribution on its own, as it is applicable to virtually any ABM for the purpose of illuminating ambiguous behavioural grounding, feeding back into social theory (generative social science), testing and scrutinising formalisations, and identifying areas for model improvement. We can be confident in the sufficiency and suitability of our analytical method. Since we followed heuristics, future work on inverse modelling should not leave uncertainty quantification unaddressed, particularly if policy considerations are to be raised.

Conclusively, has our research aim been achieved? Yes, we have studied the dynamics of EV charging behaviour by combining participatory and inverse modelling techniques. Through our new general workflow for structural-parametric calibration and interpretation of ABMs, we have gained behavioural insights, generated new research questions for behavioural psychology that yield practical policy considerations, and opened new avenues for modelling to scrutinise assumptions and implementation. Future research should consider incorporating LLM techniques along the workflow trajectory (Fig. 1) and further disentangle interrelated variables. With this present research, we have shown that for the method's broader applicability and to give inverse modelling a strong interpretative focus, it is important to do it right, to do it together.

Data availability

The data and code are available in through the URL https://doi.org/10.4121/b2c76c52-ed9a-49be-bc51-328fab4be94b.

Received: 5 February 2026; Accepted: 28 May 2026
Published online: 02 July 2026

References

1. RVO. Dutch National Charging Infrastructure Agenda. Tech. Rep., RVO (2022). File ID: 2498792.
2. van der Kam, M., van Sark, W. & Alkemade, F. Multiple roads ahead: How charging behavior can guide charging infrastructure roll-out policy. *Transp. Res. D Transp. Environ.* **85**, 102452. https://doi.org/10.1016/j.trd.2020.102452 (2020).
3. Hayn, M., Zander, A., Fichtner, W., Nickel, S. & Bertsch, V. The impact of electricity tariffs on residential demand side flexibility: results of bottom-up load profile modeling. *Energy Systems* **9**, 759–792. https://doi.org/10.1007/s12667-018-0278-8 (2018).
4. Bakhuis, J., Barbour, N., Molin, E. & Chappin, E. Understanding user preferences regarding vehicle-to-grid (V2G): A latent class choice analysis. *Transp. Res. A Policy Pract.* https://doi.org/10.1016/j.tra.2025.104610 (2025).
5. Tamis, M., De Vries, G., Renes, R. J. & Alkemade, F. Plugging in with neighbours: Defining the social dimension of electric vehicle charging in the Netherlands. *Energy Res. Soc. Sci.* **118**, 103791. https://doi.org/10.1016/j.erss.2024.103791 (2024).
6. Helmus, J. R., Lees, M. H. & Van Den Hoed, R. A data driven typology of electric vehicle user types and charging sessions. *Transp. Res. C Emerg. Technol.* **115**, 102637. https://doi.org/10.1016/j.trc.2020.102637 (2020).
7. Wolbertus, R., Kroesen, M., Van Den Hoed, R. & Chorus, C. Fully charged: An empirical study into the factors that influence connection times at EV-charging stations. *Energy Policy* **123**, 1–7. https://doi.org/10.1016/j.enpol.2018.08.030 (2018).
8. Sprei, F. & Kempton, W. Mental models guide electric vehicle charging. *Energy* **292**, 130430. https://doi.org/10.1016/j.energy.2024.130430 (2024).
9. Noel, L., Zarazua de Rubens, G., Sovacool, B. K. & Kester, J. Fear and loathing of electric vehicles: The reactionary rhetoric of range anxiety. *Energy Res. Soc. Sci.* **48**, 96–107. https://doi.org/10.1016/j.erss.2018.10.001 (2019).
10. Ashkrof, P., Homem de Almeida Correia, G. & van Arem, B. Analysis of the effect of charging needs on battery electric vehicle drivers' route choice behaviour: A case study in the Netherlands. *Transp. Res. D Transp. Environ.* **78**, 102206. https://doi.org/10.1016/j.trd.2019.102206 (2020).
11. Paradies, G. L., Usmani, O. A., Lamboo, S. & van den Brink, R. W. Falling short in 2030: Simulating battery-electric vehicle adoption behaviour in the Netherlands. *Energy Res. Soc. Sci.* **97**, 102968. https://doi.org/10.1016/j.erss.2023.102968 (2023).
12. Van Der Kam, M., Peters, A., Van Sark, W. & Alkemade, F. Agent-based modelling of charging behaviour of electric vehicle drivers. *J. Artif. Soc. Social Simul.* **22**, 7. https://doi.org/10.1564/jasss.4133 (2019).
13. Chappin, E. & Dijkema, G. P. J. Agent-based modeling of energy infrastructure transitions. In *2008 First International Conference on Infrastructure Systems and Services: Building Networks for a Brighter Future (INFRA)*, 1–6, https://doi.org/10.1109/INFRA.2008.5439580 (IEEE, Rotterdam, Netherlands, 2008).
14. Hansen, P., Liu, X. & Morrison, G. M. Agent-based modelling and socio-technical energy transitions: A systematic literature review. *Energy Res. Soc. Sci.* **49**, 41–52. https://doi.org/10.1016/j.erss.2018.10.021 (2019).
15. Chappin, E. et al. Simulating climate and energy policy with agent-based modelling: The Energy Modelling Laboratory (EMLab). *Environ. Model. Softw.* **96**, 421–431. https://doi.org/10.1016/j.envsoft.2017.07.009 (2017).
16. Collins, A., Koehler, M. & Lynch, C. Methods that support the validation of agent-based models: An overview and discussion. *J. Artif. Soc. Social Simul.* https://doi.org/10.18564/jasss.5258 (2024).
17. De Vries, G., Biely, K. & Chappin, E. Psychology: The missing link in transitions research. *Environ. Innov. Soc. Transit.* **41**, 42–44. https://doi.org/10.1016/j.eist.2021.09.015 (2021).
18. Antosz, P. et al. What do you want theory for? - A pragmatic analysis of the roles of "theory" in agent-based modelling. *Environ. Model. Softw.* **168**, 105802. https://doi.org/10.1016/j.envsoft.2023.105802 (2023).
19. Janssen, M., Alessa, L. N., Barton, C. M., Bergin, S. & Lee, A. Towards a Community Framework for Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation***11** (2008).
20. Rietkerk-van der Wijngaart, M., de Jager, L., Scholz, G., Chappin, E. & de Vries, G. Bridging behavioral theory and household energy decisions: Enhancing agent-based models with behavioral analysis. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2025.1568730 (2025).
21. Cash, P., Wrobel, A., Maier, A. & Hansen, J. P. Understanding long-term behaviour change techniques: a mixed methods study. *J. Eng. Des.* **34**, 383–410. https://doi.org/10.1080/09544828.2023.2227933 (2023).
22. Charalabidis, Y., Loukis, E. & Androutsopoulou, A. *Enhancing participative policy making through modelling and simulation: A state of the art review* (Greece, 2011).
23. Prell, C. et al. If you have a hammer everything looks like a nail: Traditional versus participatory model building. *Interdiscip. Sci. Rev.* https://doi.org/10.1179/030801807X211720 (2007).
24. Voinov, A. & Bousquet, F. Modelling with stakeholders. *Environ. Model. Softw.* **25**, 1268–1281. https://doi.org/10.1016/j.envsoft.2010.03.007 (2010).
25. Cuppen, E., Nikolic, I., Kwakkel, J. & Quist, J. Participatory multi-modelling as the creation of a boundary object ecology: The case of future energy infrastructures in the Rotterdam Port Industrial Cluster. *Sustain. Sci.* **16**(3), 901–918. https://doi.org/10.1007/s11625-020-00873-z (2021).
26. Grimm, V. & Railsback, S. F. Pattern-oriented modelling: A 'multi-scope' for predictive systems ecology. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 298–310. https://doi.org/10.1098/rstb.2011.0180 (2012).
27. Hendricks Franssen, H. J. et al. A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. *Adv. Water Resour.* **32**, 851–872. https://doi.org/10.1016/j.advwatres.2009.02.011 (2009).
28. Schubotz, L., Chappin, E. & Scholz, G. Structural Sculpting: Making Inverse Modelling Generate and Deal with Variable Structures. *J. Artif. Soc. Social Simul.* **28**, 10, https://doi.org/10.18564/jasss.5830 (2025).
29. Epstein, J. M. Inverse Generative Social Science: Backward to the Future. *J. Artif. Soc. Social Simul.* **26**, 9, https://doi.org/10.18564/jasss.5083 (2023).
30. Jensen, T. & Chappin, E. Automating agent-based modeling: Data-driven generation and application of innovation diffusion models. *Environ. Modell. Softw.* **92**, 261–268. https://doi.org/10.1016/j.envsoft.2017.02.018 (2017).
31. Greig, R., Major, C., Pacholska, M., Bending, S. & Arranz, J. Learning interpretable logic for agent-based models from domain independent primitives. *J. Artif. Soc. Social Simul.* **26**, 12. https://doi.org/10.18564/jasss.5087 (2023).

32. Gunaratne, C., Hatna, E., Epstein, J. M. & Garibay, I. Generating mixed patterns of residential segregation: An evolutionary approach. *J. Artif. Soc. Social Simul.* **26**, 7. https://doi.org/10.18564/jasss.5081 (2023).
33. Miranda, L., Garibay, O. O. & Baggio, J. Evolutionary model discovery of human behavioral factors driving decision-making in irrigation experiments. *J. Artif. Soc. Social Simul.* **26**, 11. https://doi.org/10.18564/jasss.5069 (2023).
34. Vu, T. M. et al. Can social norms explain long-term trends in alcohol use? Insights from inverse generative social science. *J. Artif. Soc. Soc. Simul.* **26**, 4. https://doi.org/10.18564/jasss.5077 (2023).
35. Udrescu, S.-M. & Tegmark, M. AI Feynman: A physics-inspired method for symbolic regression. *Sci. Adv.* **6**, eaay2631. https://doi.org/10.1126/sciadv.aay2631 (2020).
36. Senanayake, G. P. D. P. *et al.* Simultaneous Structural and Parameter Optimization in Agent-Based Models Using Adaptive Genetic Programming. *IEEE Transactions on Computational Social Systems* 1–13, https://doi.org/10.1109/TCSS.2025.3628208 (2025).
37. Epstein, J. M. Agent-based computational models and generative social science. *Complexity* **4**, 41–60. (1999).
38. Van Der Kam, M., Peters, A., Sark, W. v. & Alkemade, F. Charging behaviour of electric vehicle drivers v1.1.2, https://doi.org/10.25937/q6ey-6n41 (2020).
39. West, R. et al. Development of a formal system for representing behaviour-change theories. *Nat. Hum. Behav.* **3**, 526–536. https://doi.org/10.1038/s41562-019-0561-2 (2019).
40. Heino, M. T. J., Knittle, K., Noone, C., Hasselman, F. & Hankonen, N. Studying behaviour change mechanisms under complexity. *Behav. Sci.* **11**, 77. https://doi.org/10.3390/bs11050077 (2021).
41. Hallsworth, M. A manifesto for applying behavioural science.. *Nat. Hum. Behav.* **7**, 310–322. https://doi.org/10.1038/s41562-023-01555-3 (2023).
42. Parkinson, J. A., Gould, A., Knowles, N., West, J. & Goodman, A. M. Integrating systems thinking and behavioural science.. *Behav. Sci.* **15**, 403. https://doi.org/10.3390/bs15040403 (2025).
43. Tormala, Z. Attitudes: Form, Function, and the Factors that Shape Them. In Gilbert, D., Fiske, S., Finkel, E. & Mendes, W. (eds.) *Handbook of Social Psychlogy 6th Edition*, https://doi.org/10.70400/EOQW7740 (Situational Press, 2025), 6 edn.
44. Tajfel, H. & Turner, J. C. *An integrative theory of intergroup conflict* 33–47 (Brooks/Cole, 1979).
45. Tajfel, H. & Turner, J. C. The Social Identity Theory of Intergroup Behavior. In *Political Psychology* (eds Jost, J. T. & Sidanius, J.) 276–293 (Psychology Press, 2004). https://doi.org/10.4324/9780203505984-16.
46. Jost, J. T. & Banaji, M. R. The role of stereotyping in system-justification and the production of false consciousness.. *Br. J. Soc. Psychol.* **33**, 1–27. https://doi.org/10.1111/j.2044-8309.1994.tb01008.x (1994).
47. Jost, J. T., Banaji, M. R. & Nosek, B. A. A decade of system justification theory: Accumulated evidence of conscious and unconscious bolstering of the status quo.. *Polit. Psychol.* **25**, 881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x (2004).
48. Potoglou, D., Song, R. & Santos, G. Public charging choices of electric vehicle users: A review and conceptual framework.. *Transp. Res. D. Transp. Environ.* **121**, 103824. https://doi.org/10.1016/j.trd.2023.103824 (2023).
49. Pevec, D. et al. A survey-based assessment of how existing and potential electric vehicle owners perceive range anxiety.. *J. Clean. Prod.* **276**, 122779. https://doi.org/10.1016/j.jclepro.2020.122779 (2020).
50. Wolbertus, R. & Gerzon, B. Improving electric vehicle charging station efficiency through pricing.. *J. Adv. Transp.* **2018**, 4831951. https://doi.org/10.1155/2018/4831951 (2018).
51. Ozili, P. K. The Acceptable R-Square in Empirical Modelling for. *Soc. Sci. Res.* https://doi.org/10.2139/ssrn.4128165 (2022).
52. Michie, S., van Stralen, M. M. & West, R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement. Sci.* **6**, 42. https://doi.org/10.1186/1748-5908-6-42 (2011).
53. Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. User acceptance of information technology: Toward a unified view. *MIS Q.* **27**, 425–478. https://doi.org/10.2307/30036540 (2003).
54. Kahneman, D. & Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **47**, 263–291. https://doi.org/10.2307/1914185 (1979).
55. Dablander, F., Lange, F., Brick, C. & Aron, A. R. Expressing intentions is not climate action. *Proc. Natl. Acad. Sci.* **122**, e2512457122. https://doi.org/10.1073/pnas.2512457122 (2025).
56. Urai, A. E., Dablander, F. & Bolderdijk, J. W. Impactful environmental psychology needs formal theories. *Curr. Opin. Behav. Sci.* **66**, 101619. https://doi.org/10.1016/j.cobeha.2025.101619 (2025).
57. Egbue, O., Long, S. & Samaranayake, V. A. Mass deployment of sustainable transportation: evaluation of factors that influence electric vehicle adoption. *Clean Technol. Environ. Policy* **19**, 1927–1939. https://doi.org/10.1007/s10098-017-1375-4 (2017).
58. Tamis, M., de Vries, G. & Renes, R. J. Moving your electric car for others: Identifying injunctive norms for sharing public charging points. *Travel Behaviour and Society* **44**, 101290. https://doi.org/10.1016/j.tbs.2026.101290 (2026).
59. Wijermans, N. *et al.* Models in Social Psychology and Agent-Based Social simulation – an interdisciplinary conversation on similarities and differences. *Review of Artificial Societies and Social Simulation***10** (2022).
60. Yu, J. & McKinley, G. Synthetic participatory planning of shared automated electric mobility systems. *Sustainability* **16**, 5618. https://doi.org/10.3390/su16135618 (2024).
61. van de Schoot, R. et al. An open source machine learning framework for efficient and transparent systematic reviews. *Nat. Mach. Intell.* **3**, 125–133. https://doi.org/10.1038/s42256-020-00287-7 (2021).
62. Liang, Y. et al. Analyzing sequential activity and travel decisions with interpretable deep inverse reinforcement learning. *Travel Behav. Soc.* **43**, 101171. https://doi.org/10.1016/j.tbs.2025.101171 (2026).
63. Senanayake, G. P. D. P. & Kieu, M. Inverse Generative Approach for Identifying Agent-Based Models from Stochastic Primitives. In *Multi-Agent-Based Simulation XXV* (eds Thompson, J. & Stankov, I.) 70–86 (Springer Nature Switzerland, 2025). https://doi.org/10.1007/978-3-031-88017-9_6.
64. Senanayake, G. P. D. P. & Kieu, M. An adaptive evolutionary approach for discovering stochastic agent-based models. *J. Artif. Soc. Soc. Simul.* **28**, 1. https://doi.org/10.18564/jasss.5737 (2025).
65. Herman, J. & Usher, W. SALib: An open-source Python library for sensitivity analysis. *J. Open Source Softw.* **2**, 97. https://doi.org/10.21105/joss.00097 (2017).
66. Iwanaga, T., Usher, W. & Herman, J. Toward SALib 2.0: Advancing the accessibility and interpretability of global sensitivity analyses. *Socio-Environ. Syst. Model.* **4**, 18155–18155. https://doi.org/10.18174/sesmo.18155 (2022).
67. Lennon, M. & Rabbani, J. Driving equity: On the extractive ecosystem of electric vehicles. *Environ. Soc.* **16**, 69–89. https://doi.org/10.3167/ares.2025.160105 (2025).
68. DeWit, A. Decarbonization and Critical Raw Materials. In *Governance for a Sustainable Future: The State of the Art in Japan* (eds Adachi, Y. & Usami, M.) 279–298 (Springer Nature, 2023). https://doi.org/10.1007/978-981-99-4771-3_14.
69. Jaxa-Rozen, M. & Kwakkel, J. H. PyNetLogo: Linking NetLogo with Python. *J. Artif. Soc. Soc. Simul.* **21**, 4. https://doi.org/10.18564/jasss.3668 (2018).
70. Wilensky, U. *Netlogo* (Center for Connected Learning and Computer-Based Modeling, 1999).
71. Tavenard, R. An introduction to Dynamic Time Warping (2021).
72. Filatova, T. *et al.* AGENTBLOCKS: a community platform for sharing, comparing, and improving reusable building blocks for (agent-based) models, https://doi.org/10.2139/ssrn.5219387 (2025).
73. Berger, U. et al. Towards reusable building blocks for agent-based modelling and theory development. *Environ. Model. Softw.* **175**, 106003. https://doi.org/10.1016/j.envsoft.2024.106003 (2024).
74. Tavenard, R. *et al.* Tslearn, A Machine Learning Toolkit for Time Series Data. *Journal of Machine Learning Research***21**, 1–6 (2020). TSLearn Website: https://tslearn.readthedocs.io/en/latest/index.html

75. Sakoe, H. & Chiba, S. Dynamic programming algorithm optimization for spoken word recognition. *IEEE Trans. Acoust. Speech Signal Process.* **26**, 43–49. https://doi.org/10.1109/TASSP.1978.1163055 (1978).
76. Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustration in the case of the binomial. *Biometrika* **26**, 404–413. https://doi.org/10.1093/biomet/26.4.404 (1934).
77. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794, https://doi.org/10.1145/2939672.2939785 (ACM, San Francisco California USA, 2016).
78. Lundberg, S. M. & Lee, S.-I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems*, vol. 30, https://doi.org/10.48550/arXiv.1705.07874 (Curran Associates, Inc., 2017).
79. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodological)* **57**, 289–300 (1995).
80. Haynes, W. Benjamini–Hochberg Method. In *Encyclopedia of Systems Biology*, 78–78, https://doi.org/10.1007/978-1-4419-9863-7_1215 (Springer, New York, NY, 2013).
81. DHPC. *DelftBlue Supercomputer (Phase 2)* (Delft High Performance Computing Centre, 2024).
82. Kazil, J., Masad, D. & Crooks, A. Utilizing Python for Agent-Based Modeling: The Mesa Framework. In *Social, Cultural, and Behavioral Modeling* (eds Thomson, R. et al.) 308–317 (Springer International Publishing, 2020). https://doi.org/10.1007/978-3-030-61255-9_30.
83. Egner, L. E. Household Energy Retrofit Behavior (HERB) ABM (2022) (v1.0.0).
84. Boria, E. Neighbor Influenced Energy Retrofit (NIER) agent-based model (2020) (v1.0.0).

Acknowledgements
This research has been made possible through funding by the Energy Transition Lab of TU Delft's Faculty of Technology, Policy and Management. We thank the TU Delft Library for covering the article processing charges. The authors thank a friendly reviewer who, in the early stages of the manuscript, believed that this was "fully deserving of the compliment of rational opposition".

Author contributions
Conceptualisation and design: all authors. Modelling, interviews, data analysis, figures, writing, reviews: L.S. E.C. and G.S. and G.d.V. supervised and helped with conceptualisation and interpretation. All authors reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Ethics approval
Ethical approval that this study was conducted and all methods carried out in accordance with relevant guidelines and regulations was obtained from Human Research Ethics Committee TU Delft (approval number 6380) and informed consent given by all participants.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process
During the preparation of this work, the authors used the following tools for the following reasons: ChatGPT to identify unclear articulations, and GitHub Copilot to debug and accelerate code. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Positioning statement
It is important to emphasise that the goal of this research is to illuminate behavioural dynamics, not to advocate for or against EVs. Like other technologies promoted under the umbrella of "decarbonisation", EVs depend on raw materials and rare earth metals and, therefore, raise justice concerns regarding access to these, environmental degradation, and human health harms[67,68]. We encourage readers to understand EVs not as a neutral instrument in a local transition but embedded in broader environmental dilemmas and socio-political complexities.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-026-56059-7.

Correspondence and requests for materials should be addressed to L.S.