SILESIAN UNIVERSITY OF TECHNOLOGY PUBLISHING HOUSE

SCIENTIFIC PAPERS OF SILESIAN UNIVERSITY OF TECHNOLOGY 2026
ORGANIZATION AND MANAGEMENT SERIES NO. 245

EPIDEMIC SEIR MODEL OF THE E-COMMERCE CUSTOMER LIFE-CYCLE FOR DEMAND ESTIMATION UNDER TIME-VARYING MARKETING EXPENDITURES AND SEASONAL EFFECTS

Marek MĘDREK

Maria Curie-Sklodowska University, Faculty of Economics, Department of Information Systems and Logistics; marek.medrek@umcs.pl, ORCID: 0000-0001-5752-5084

Purpose: This paper develops a weekly e-commerce demand model that explicitly captures exposure, activation, churn, and reactivation in the customer life-cycle during purchasing process dynamics under time-varying marketing pressure and seasonality.
Design/methodology/approach: Using three years of weekly transaction data from an online retailer, along with real marketing expenses and seasonal effects, we calibrate a latent-state-space SEIRS model of the customer life-cycle in the purchasing process. We introduce a time-varying transmission parameter that aggregates marketing pressure, social influence, and seasonality, while separate transitions between customer states capture customer churn and reactivation.
Findings: The SEIRS model outperforms the SARIMA benchmark in test-window forecasts and provides interpretable diagnostics, including transition paths between customer states and a demand indicator derived from the difference between active and outgoing customers. Scenario analysis shows that different customer churn and reactivation rates generate different demand trajectories.
Research limitations/implications: Models of individual state spaces, inspired by epidemiology, can be used for marketing purposes by separating exposure, conversion, and flow processes across states in the life cycle, offering more diagnostic possibilities than standard time-series models.
Practical implications: The model's parameters reflect factors influencing demand and conversion dynamics in the purchasing process. Monitoring the transmission parameter λ_t, the momentum proxy R_t, and the churn and reactivation parameters (γ, ω) support coordination of marketing expenditures and seasonal revenues. For managers for whom demand forecasting is a key issue, it may also be helpful to identify the customer life-cycle mechanisms that drive demand changes: customer acquisition, customer churn, and customer reactivation. Without understanding these mechanisms, general principles form the basis for campaign budgets, which risks overspending during periods of low demand and underinvestment when conditions are favorable for recovery.
Originality/value: The paper proposes a compact life-cycle specification of the customer life-cycle in the purchasing process, that treats churn and reactivation as explicit, estimable state transitions in a weekly demand system, combining strong forecasting with operational interpretability.

Keywords: Marketing analytics; E-commerce demand; State-space modeling; Decision support.
Category of the paper: Research paper.

1. Introduction

E-commerce demand is rarely stable. Many retailers see sudden surges followed by stable periods, slowdowns, and renewed activity, driven by social influence, marketing, and calendar-based seasonality. From a marketing analytics view, this reflects a recurring customer life-cycle: consumers become aware and active, some churn, and many later return through reactivation rather than only through first-time adoption. Yet classical diffusion models, such as the Bass model (Bass, 1969), focus on cumulative adoption of new products via innovators and imitators and do not explicitly model exposure, dropouts, or reactivation (Peres et al., 2010; Rogers, 2003). This limits their value for managers who need measurable signals to predict customer churn and plan reactivation. In practice, churn analytics is often framed as a predictive classification task, but recent work highlights the importance of time-to-event structure and explainability when churn signals guide the timing of retention interventions (Dias, Antonio, 2025; Kimitei et al., 2025).

To capture these stages in the customer life-cycle in e-commerce processes, we use compartmental diffusion models (SIR/SEIR/SEIRS) that explicitly represent exposure, delayed activation, loss of activity, and possible market reentry (Kermack, McKendrick, 1927; Medrek, Pastuszak, 2021; Tillett, 1992). Originally developed for epidemiology, these models are now widely applied to digital behavior and adoption, where awareness and engagement spread via social contact and information exchange (Fibich, 2016a; McAdams, Song, 2025). Related work proposes hybrid diffusion–epidemic models, such as Bass–SIR variants, to capture periods of limited customer influence and subsequent inactivity (Peres et al., 2010). Empirical and simulation studies in online settings likewise show that the path from exposure to activity can mirror the transition from awareness to purchase in e-commerce (Woo, Chen, 2016). These findings point to models that track exposure, activation, churn, and reactivation of customers as distinct stages rather than components of fixed curves. Customer analytics is increasingly used to anticipate churn and proactively optimise commercial actions (Wolniak, 2024), aligning with the use of latent-state dynamics as early-warning signals for intervention timing.

Although customer marketing analytics applications are very promising, significant challenges remain. One such example is the difficulty of accurately attributing changes in demand to internal feedback or external marketing pressure, as both evolve and their impacts are interconnected. Additionally, using a reproduction number outside of epidemiology is controversial. If used, it must be consistently defined and estimated from available data to avoid

misleading interpretations (Nash et al., 2022). These issues require a framework that links time-dependent demand to marketing and calendar inputs while keeping customer life-cycle stages (e.g., churn and reactivation) explicit and estimable.

This paper models weekly e-commerce demand using a latent-state SEIRS framework that explicitly represents stages in the customer life-cycle, i.e., exposure, activation, churn, and reactivation under time-varying transmission between stages, driven by social, marketing, and calendar effects. By linking latent dynamics to real store data, the method gives metrics ($\lambda_t, \gamma, \omega$) and a reproduction proxy that we use as a demand signal. We test the model's validity by comparing conversion forecasts to a seasonal time-series benchmark SARIMA. Our work is both theoretical and practical. We demonstrate the life-cycle forces driving demand and provide a tool for managing seasonal demand through marketing pressure.

We test the proposed approach through four research hypotheses.

H1: The epidemic customer life cycle model improves e-commerce demand forecasts compared to the standard benchmark.

H2: The SEIR model provides early warning signals in e-commerce management.

H3: Customer loss and reactivation mechanisms influence demand trajectories.

H4: The SEIRS model provides an interpretable indicator, akin to the epidemic reproduction number, that identifies expansive and recessive configurations of e-commerce systems in terms of order numbers.

The paper is organised as follows. Section 2 provides a literature review, discusses new product adoption, diffusion, and compartmental approaches, and motivates the extension of the customer reactivation stage through modelling. Section 3 presents the model specification, calibration, and case data. Section 4 reports forecasting performance, diagnostics, and sensitivity scenario analyses, and Section 5 concludes with managerial implications, limitations, and directions for further work.

2. Literature review: customer life-cycle dynamics in e-commerce demand

In this part of our work, we place the proposed SEIRS model of customer life cycle states in the e-commerce purchasing process within two strands of marketing analytics research: diffusion models explaining long-term penetration (people who always use a given service) and offer contagion models based on stages of exposure, customer activation, and influence decay. The primary motivation is that demand in e-commerce is shaped not only by one-time acceptance of an offer, but also by the recurring dynamics of the customer life cycle: temporary inactivity (churn) and return (reactivation/recovery). Section 2.1 discusses Bass's model and explains why cumulative penetration is insufficient for modeling real demand. Section 2.2 summarizes SEIR network diffusion models and their ability to represent stage-based activation

with limited impact. Section 2.3 identifies a gap in SEIR models for e-commerce demand and justifies a minimal extension of SEIRS models to include a customer reactivation to the initial stage. The final section (Section 2.4) presents the design requirements for state space models such as the proposed SEIRS model.

2.1. Bass diffusion for penetration

The Bass model remains a reliable tool for modeling the timing of the initial adoption of new products. It assumes that the probability of purchase at time t is linearly related to the cumulative number of previous adopters, capturing two mechanisms: innovation (external influence) and imitation (internal influence) (Bass, 1969). This behavioral premise yields an adoption-rate curve that rises to a single peak and then decays as the market saturates.

Formally, the adoption rate $f(t)$ and the cumulative share of adopters $F(t)$ follow

$$f(t) = \frac{dF(t)}{dt} = (p + qF(t))[1 - F(t)] \tag{1}$$

and

$$F(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}} \tag{2}$$

where p is the coefficient of innovation and q the coefficient of imitation (Bass, 1969). Figure 1 illustrates the characteristic S-shaped cumulative diffusion and its corresponding adoption-rate peak.

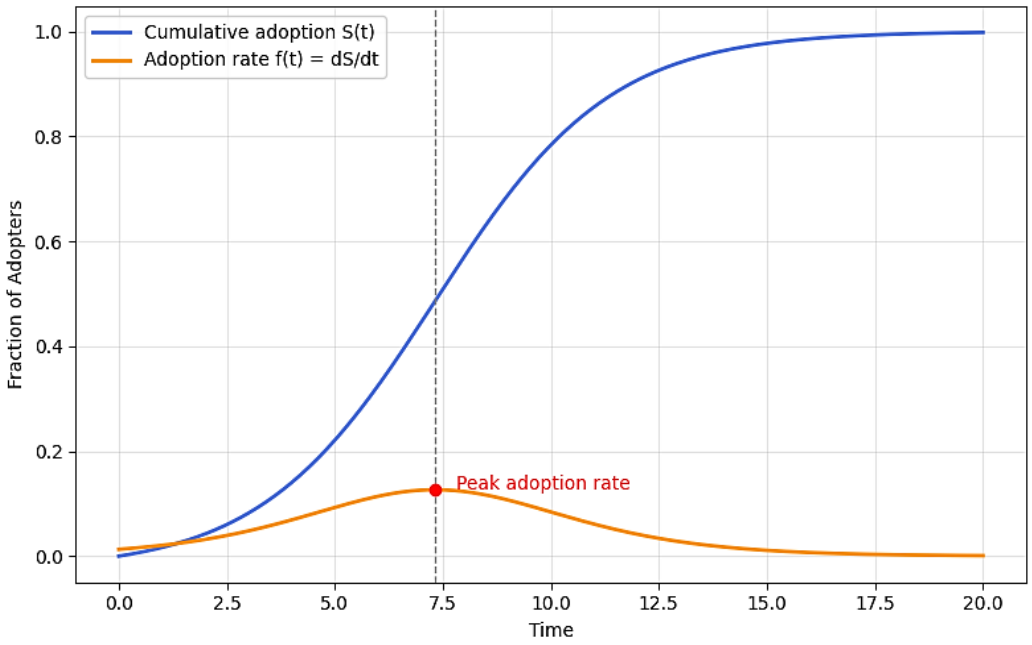

Figure 1. Traditional Bass diffusion model (Bass, 1969). The blue line S(t) is the cumulative adoption share, forming an S-curve that starts slowly, speeds up through imitation, and then nears saturation. The orange curve f(t) = dS/dt is the adoption rate, peaking when innovation and imitation balance; the red point marks the maximum adoption intensity.

Although empirically successful in modeling innovation penetration, the Bass model is structurally misaligned with recurring demand in digital retail. It assumes a homogeneous population, ignores the structure of the social network and its heterogeneous influence, and treats market potential as fixed, with a one-way transition to adoption, which is reasonable for "ever-adopters" but not for actual consumer behavior, where customers cycle in and out. The model also omits repeat purchases and reactivation, making it unsuitable for the cyclical buying and return patterns typical of modern e-commerce. Finally, it does not explicitly capture external shocks, such as marketing expenditures, social influence (media word of mouth effect), or macroeconomic changes, that can strongly affect diffusion in real digital markets (Peres et al., 2010; Rogers, 2003). In terms of marketing analytics, Bass explains the cumulative penetration[1] rather than the active customer base that generates weekly orders.

2.2. SEIR models for staged adoption

To model staged behavioral dynamics, diffusion studies in social systems increasingly use SIR/SEIR compartmental frameworks. In these approaches, consumers move from susceptible (S; reachable but not yet influenced), through exposed (E; aware/considering) and active/infectious (I; purchasing and potentially influencing others), and finally to removed (R; no longer influencing) (Kermack, McKendrick, 1927; Medrek, Pastuszak, 2021b; Tillett, 1992). In discrete time, a standard SEIR update in a closed population $N = S_t + E_t + I_t + R_t$ can be written as

$$\begin{aligned} S_{t+1} &= S_t - \beta \frac{S_t I_t}{N}, \\ E_{t+1} &= E_t + \beta \frac{S_t I_t}{N} - \sigma E_t, \\ I_{t+1} &= I_t + \sigma E_t - \gamma I_t, \\ R_{t+1} &= R_t + \gamma I_t, \end{aligned} \tag{3}$$

where β is the transmission intensity, σ the exposure to activation conversion, and γ the rate at which influence/activity decays.

A key advantage of such interpretations is that transmission depends on who interacts with whom: exposure risk rises with neighbour activity, and diffusion speed becomes sensitive to local structure rather than a single aggregate path (Prashad, 2025). This aligns with marketing evidence that awareness and adoption spread through word of mouth (WOM), information exchange, and belief updating (McAdams, Song, 2025; Woo, Chen, 2016). Hybrid diffusion–contagion models (e.g., Bass–SIR variants) formalise finite influence windows and subsequent inactivity, improving realism of the modelled life-cycle when attention decays quickly (Peres et al., 2010). SEIR-type models further separate exposure formation, delayed conversion,

[1] Penetration is the number of unique users who have adopted at least once; it does not necessarily reflect current activity.

and influence decay, and naturally model time-varying transmission to reflect a changing market context.

2.3. Customer churn and reactivation as fundamental modeling primitives: an SEIRS-based extension

Repeated purchases and recurring engagement drive weekly e-commerce demand. Customers cycle through awareness and purchase, then become inactive, and later return via seasonal revisits, campaign-driven comebacks, or renewed interest after assortment or price changes. Classical diffusion models emphasize long-term penetration and omit churn and reactivation, while standard SEIR models assume removed users never return, which is inconsistent with the dynamics of customer recovery (Fibich, 2016b; Nash et al., 2022; Papageorgiou, Tsaklidis, 2023). Modeling these dynamics requires an explicit loop from the inactivity stage back to the susceptibility stage.

We therefore adopt the SEIR extension that introduces reactivation, i.e., SEIRS. In discrete time, we interpret compartments as: Susceptible (S_t; able to buy but not purchasing this week), Exposed (E_t; aware/considering), Active (I_t; purchased this week), and Removed (R_t; temporarily inactive). Reactivation is implemented as a return flow $R \to S$ governed by $\omega \in (0,1)$:

$$\begin{aligned} S_{t+1} &= S_t - \lambda_t S_t + \omega R_t, \\ E_{t+1} &= E_t + \lambda_t S_t - \sigma E_t, \\ I_{t+1} &= I_t + \sigma E_t - \gamma I_t, \\ R_{t+1} &= R_t + \gamma I_t - \omega R_t \end{aligned} \tag{4}$$

where λ_t is time-dependent transmission intensity ($S \to E$), σ is conversion ($E \to I$), and γ is churn ($I \to R$). Figure 2 summarizes the resulting looped life-cycle mechanism.

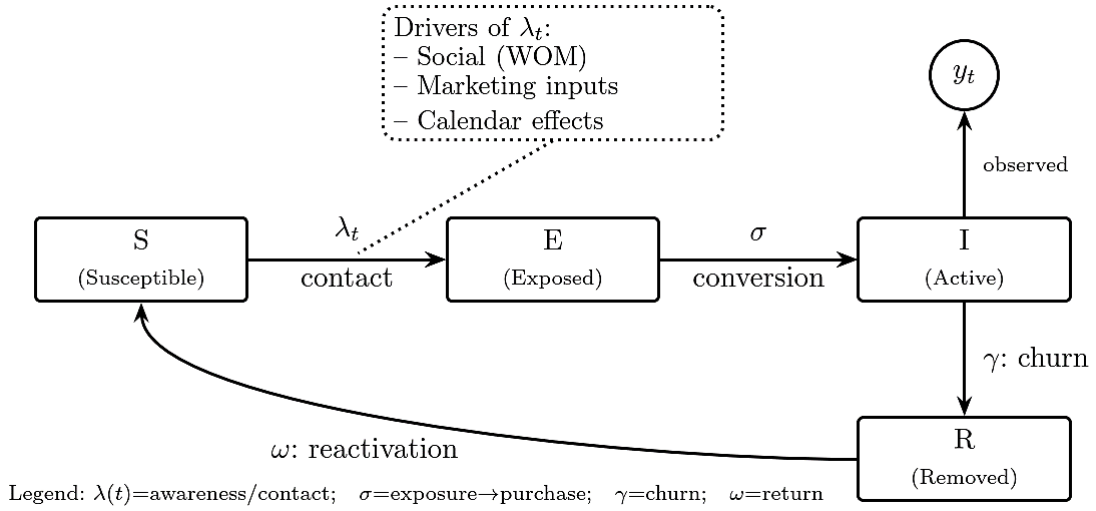

Figure 2. SEIRS transition diagram for e-commerce with observed output y_t and drivers λ_t. Arrows denote flows between states S, E, I, R, S in time step t.

In marketing analytics, it is crucial that the parameter λ_t must vary over time and aggregate multiple factors. The social diffusion of demand depends on the structure of the network and the decay of attention (McAdams, Song, 2025; Prashad, 2025), the belief updates shape the demand in the adoption epidemic models (Woo, Chen, 2016), and in digital commerce the effects of paid media and calendars can influence the exposure to offers and the intentions of potential customers. We therefore model λ_t as a time-varying intensity driven by internal WOM/social activity, external marketing inputs, and calendar/exogenous shocks. The reactivation loop offers two practical benefits. Primarily, it explains repeated waves of customer activity: peaks can be driven by increased conversion (σ) and/or by a larger customer reactivation inflow (ω) that replenishes the susceptible pool available for conversion. It also provides a breakdown of demand dynamics into transmission and departures of customers, and the SEIRS structure allows for a consistent mapping between estimated values of λ_t, γ, and the persistence of customer activity.

2.4. Design requirements for an interpretable state-space demand model

The literature identifies three design needs for an e-commerce demand model that supports marketing decisions. Initially, the model should go beyond total reach to incorporate stage trends (exposure and activation) and life-cycle flows (churn and reactivation). Second, transmission must be time-varying and interpretable, distinguishing endogenous social effects from exogenous marketing and calendar drivers (Rogers, 2003). Lastly, any reproduction-style indicator must be derived from the same model used to estimate transmission and be consistent with the measurement scale (Nash et al., 2022; Papageorgiou, Tsaklidis, 2023).

Guided by these requirements, we operationalise a state-space SEIRS model that treats exposure, churn, and customer reactivation as state transitions between stages of the customer life-cycle. We link the latent states to store-level measurements and optimise the model's parameters using weekly observations. The state equations describe the latent dynamics and map activity to orders. The model also integrates weekly observations into forecasting: it explains why forecasts deviate around turning points by attributing changes to transmission (marketing/social/calendar), churn, and reactivation mechanisms. Thus, the approach extends diffusion-style thinking toward repeat purchase and win-back dynamics, which are central to e-commerce demand management.

3. Methods

In this section, we start with the model specification, including state dynamics, transmission, the observation equation, and a demand-momentum indicator. Then, we describe calibration and initialization, followed by the case data source and construction of marketing and seasonal drivers.

3.1. Numerical model

3.1.1. State variables and transition dynamics

We model weekly customer dynamics with a latent state vector $X_t = (S_t, E_t, I_t, R_t)$, partitioning addressable customers into Susceptible (S_t, able to buy but not purchasing this week), Exposed (E_t, aware or considering), Active (I_t, at least one purchase this week), and Removed (R_t, temporarily inactive). In the closed-population case,

$$N = S_t + E_t + I_t + R_t = \text{const}, \tag{5}$$

with all non-zero components, $S_t, E_t, I_t, R_t > 0$.

To accommodate slow changes in the reachable base (e.g., net sign-ups, market expansion), we allow an external net flow b_t into S_t. The weekly SEIRS updates become

$$\begin{aligned} S_{t+1} &= S_t - \lambda_t S_t + \omega R_t + b_t \\ E_{t+1} &= E_t + \lambda_t S_t - \sigma E_t \\ I_{t+1} &= I_t + \sigma E_t - \gamma I_t \\ R_{t+1} &= R_t + \gamma I_t - \omega R_t, \end{aligned} \tag{6}$$

and the total population evolves as

$$N_{t+1} = N_t + b_t. \tag{7}$$

We restrict b_t to be small relative to N_t to avoid confounding slow base drift with the exposure process governed by λ_t. Transition rates are weekly probabilities: churn $I \to R$ is $\gamma \in (0,1)$ and reactivation $R \to S$ is $\omega \in (0,1)$.

3.1.2. Time-varying transmission with marketing and calendar drivers

To reflect social influence, paid marketing inputs, and calendar/external effects, we model the transmission intensity as a bounded, time-varying quantity:

$$\lambda_t = \frac{1}{1 + \exp(-\eta_t)}, \tag{8}$$

$$\eta_t = \beta_0 + \beta_1 \left[\left(\frac{I_t}{N_t}\right)^{\alpha} - \overline{\left(\frac{I}{N}\right)^{\alpha}} \right] + \delta \widetilde{M}_t + \theta Z_t, \tag{9}$$

where β_0 captures the baseline exposure tendency, and the centered term $(\frac{I_t}{N_t})^{\alpha} - \overline{(\frac{I}{N})^{\alpha}}$ captures endogenous social dependence: higher activity shares increase the chance that susceptible individuals are exposed. The bar denotes the training window mean of $(\frac{I_t}{N_t})^{\alpha}$, so that β_0 retains the interpretation of typical baseline exposure when deviations in social activity, marketing, and seasonality are absent.

The exponent $\alpha > 0$ allows a flexible nonlinear response:

$$\alpha = \begin{cases} 1 & \text{linear in the active share,} \\ > 1 & \text{threshold-like (weak at low activity, strong once activity is high),} \\ < 1 & \text{diminishing returns (strong early, then saturates).} \end{cases} \quad (10)$$

Marketing influence is represented by $\boldsymbol{\delta}\widetilde{M}_t$, where M_t may include marketing expenses (e.g., email volume or paid promotion). Positive values increase exposure, while negative weights reflect poor targeting or adverse effects. Calendar/external effects (holidays, paydays, seasonality proxies, logistics shocks) enter via $\boldsymbol{\theta}Z_t$. All inputs are centered so that β_0 still corresponds to the typical weekly baseline when marketing and seasonal deviations are absent.

3.1.3. Observation equation and demand-momentum indicator

The observed weekly e-commerce outcome (orders, active buyers, or revenue) is linked to the latent active state through

$$y_t = \kappa \cdot I_t + \epsilon_t, \quad (11)$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and κ is the outcome scaling parameter (see Table 1).

We define a reproduction style demand-momentum proxy as

$$\Re_t = \frac{\lambda_t}{\gamma}, \quad (12)$$

where λ_t is the intensity of the weekly transformation $S \to E$, and γ is the weekly churn rate $I \to R$. Values $\Re_t > 1$ indicate self-sustaining demand in the absence of additional lift, while $\Re_t < 1$ signals decay unless marketing or seasonality increases the transmission intensity λ_t.

3.2. Model calibration

We calibrate the model using weekly retailer data: target y_t (number of orders), marketing input M_t (paid exposure proxy), and calendar/seasonality driver Z_t. The time step equals one week. The parameter vector is

$$\Theta = \{\beta_0, \beta_1, \alpha, \boldsymbol{\delta}, \boldsymbol{\theta}, \sigma, \gamma, \omega, \kappa, \sigma_\epsilon\}. \quad (13)$$

Latent states follow

$$X_t = (S_t, E_t, I_t, R_t), \quad (14)$$

with mass balance given by eq.(5). We set the initial reachable population to N_0, corresponding to the number of customers that can be reached at t_0. In our case, N_0 is set to the 52-week unique visitors to the store website, which is observable and sufficiently stable for weekly

dynamics, and consistent with the observation equation (11) where κ absorbs purchase intensity. To reflect the increase in exposure, we set the initial value of E_0 to small part of N_0. Then, based on the initial observation, the activity I_0 is determined. Similarly, we can infer recent inactivity and close the mass as follows

$$\begin{aligned} E_0 &= \varepsilon_E \cdot N_0, (\varepsilon_E << 1), \\ I_0 &\approx \frac{y_0}{\kappa}, \\ R_0 &= \rho_R \cdot N_0, \\ S_0 &= N_0 - E_0 - I_0 - R_0, \end{aligned} \tag{15}$$

where ρ_R reflects the initial share of lapsed but reachable customers.

Initial values of transition rates σ, γ, ω are set within economically plausible ranges (e.g., σ^{-1} about 1–2 weeks, γ^{-1} a few weeks, and ω^{-1} aligned with seasonal revisit horizons), and the final estimates are obtained by minimizing RMSE on the training window with boundary constraints $0 < \sigma, \gamma, \omega < 1$. Transmission parameters $(\beta_0, \beta_1, \alpha, \boldsymbol{\delta}, \boldsymbol{\theta})$ are estimated via eqs. (8)-(9) to align λ_t with observed turning points while avoiding overfitting of state dynamics. The scale parameter κ is estimated from the regression of y_t on I_t for the selected λ_t, and σ_ϵ is set to the standard deviation of the residuals.

Parameter estimation was performed using constrained, gradient-based optimisation on the training window, with initial conditions selected to ensure stable state trajectories. To support reproducibility, the complete parametrisation (bounds, initialisation, and driver transformations) can be provided upon reasonable request; the underlying retailer-level dataset is proprietary.

3.3. Case data and driver construction

The empirical analysis uses weekly transaction and exposure data from a medium-sized Polish online retailer specialising in gardening and construction tools. The dataset covers three years (from early 2023) and includes weekly orders y_t, paid marketing expenditure M_t (Google Ads), and calendar-based seasonal indicators Z_t.

Figure 3 shows the weekly orders series used for training (grey) and testing (black). Peaks correspond to periods of intensified paid exposure and/or seasonal triggers.

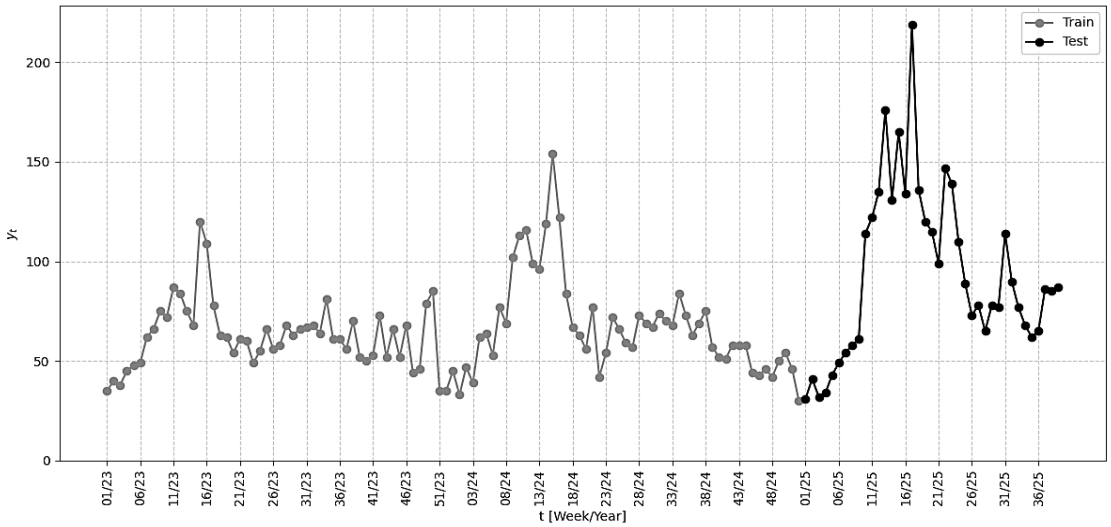

Figure 3. Weekly observed number of orders y_t representing active demand dynamics used for SEIRS model training and calibration (gray line) and testing (black line).

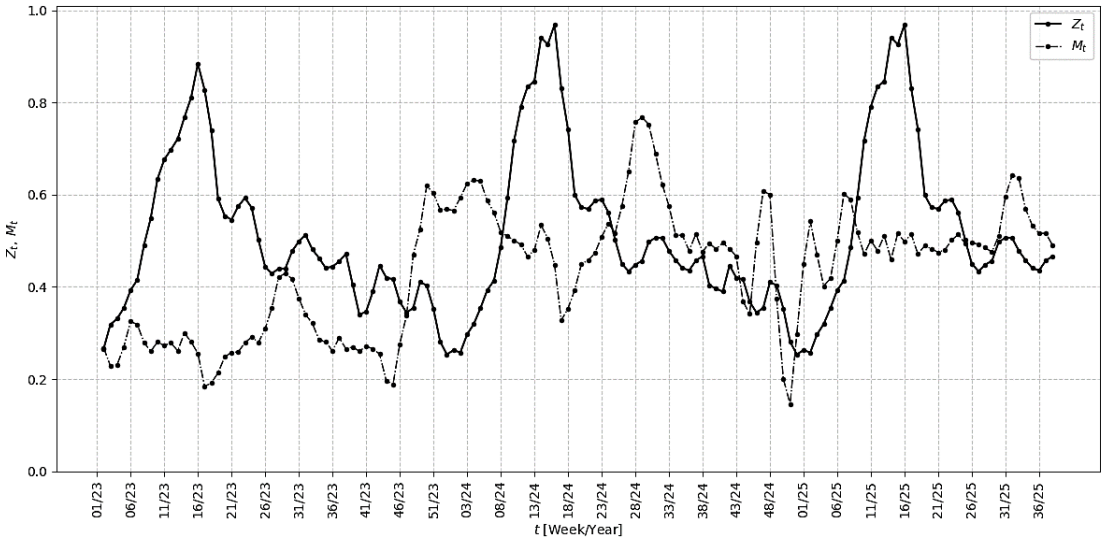

Figure 4. Time-varying marketing M_t and seasonal drivers Z_t of transmission intensity λ_t. Normalized Google Ads spend M_t is denoted by dash-dot line, and the seasonal calendar effect Z_t is marked with a solid line.

Figure 4 shows the time-varying marketing and seasonal drivers used to modulate λ_t in eqs. (8-9). Seasonality is estimated using a ratio-to-trend approach: a smooth trend T_t is obtained via a symmetric moving average ($n_t = 16$ weeks), and a raw seasonal multiplier is computed as $s_t = y_t/T_t$. To include seasonality additively in the linear predictor, we center the seasonal index:

$$Z_t = s_t - \bar{s}, \qquad \bar{s} = \frac{1}{n}\sum_{t=1}^{n_t} s_t. \tag{16}$$

The marketing driver, whose value corresponds to weekly marketing expenditures, is centered analogously, $\widetilde{M}_t = M_t - \bar{M}$, so that β_0 represents the baseline exposure tendency when marketing and seasonal deviations are absent. Such centering stabilizes identification when decomposing transmission into social, marketing, and calendar components.

Table 1 reports the final estimates of the parameter vector Θ defined in eq. (13). All external drivers (marketing and seasonality) were first transformed (smoothed over time if necessary). Parameters were then fitted on the training window using a gradient-based procedure with boundary constraints, minimizing RMSE.

With centered drivers, the baseline of the linear predictor η_t is set by the intercept $\beta_0 = -0.817$, implying a baseline weekly transmission of

$$\lambda = \text{logit}^{-1}(\beta_0) = \text{logit}^{-1}(-0.817) = 0.306. \quad (17)$$

The social component is strong but operates on a small-scale driver: $\beta_1 = 9.255$ and $\alpha = 0.960$ imply that $\beta_1(I_t/N)^\alpha$ increases rapidly even at low activity shares, generating sizeable changes in η_t (and thus λ_t) when I_t/N rises from about 1-3% to 5-8%. Marketing and seasonality inputs enter through centered (and, where appropriate, standardised) regressors, which preserves an interpretation of β_0 as the typical baseline exposure tendency.

Table 1.
Calibrated parameters, components of Θ described by eq.(13)

Symbol	Interpretation	Estimate
β_0	Baseline exposure tendency in the linear predictor η_t	$-8.1720e-01$
β_1	Strength of social influence on $(I_t/N)^\alpha$	$9.2548e+00$
α	Nonlinearity of social influence	$9.6000e-01$
$\boldsymbol{\delta}$	Weights on effective marketing input $\mathbf{M}_{nM}$	$6.9000e-01$
$\boldsymbol{\theta}$	Weights on seasonal/exogenous drivers $\mathbf{Z}_{nZ}$	$3.4400e+00$
σ	Exposure→active conversion rate ($E \to I$) per week	$8.0000e-01$
γ	Churn ($I \to R$) per week	$4.6100e-01$
ω	Reactivation ($R \to S$) per week	$2.6700e-01$
κ	Per-active output scaling in $y_t = \kappa I_t + \varepsilon_t$	$1.0000e+02$
σ_ε	Observation noise standard deviation in $y_t = \kappa I_t + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2)$	$1.4000e+01$

Transition rates are weekly: $\sigma = 0.80$ implies an average exposure to active time of $1/\sigma \approx 1.25$ weeks, $\gamma = 0.461$ implies an average active duration of $1/\gamma \approx 2.17$ weeks, and $\omega = 0.267$ implies an average return to susceptible time of $1/\omega \approx 3.75$ weeks, consistent with monthly revisit patterns.

4. Results, diagnostics, and comparative performance

This section evaluates the empirical performance of the state-space SEIRS model using a test window not used during training. We present forecasts, residual diagnostics, and behavioral interpretability via inferred latent states and dynamic indicators. These include

the time-varying transmission intensity λ_t and the demand-momentum proxy $\Re_t$. Additionally, we summarize a sensitivity analysis to assess the impact of uncertainty in key parameters on forecasting accuracy. Overall, our objective is to demonstrate that the proposed life-cycle representation enhances short-term prediction and provides a transparent decomposition of demand into exposure, activation, churn, and reactivation mechanisms that managers can interpret.

4.1. Forecast accuracy SEIRS vs. SARIMA

Figure 5 compares weekly orders y_t predicted by the SEIRS model (red) and a SARIMA benchmark (blue) against the observed number of orders (black). The vertical line marks the start of the test window (the beginning of 2025). For readability, the plots in this section are restricted to the last two years of the sample. The figure shows that SEIRS tracks turning points earlier and reproduces peak amplitudes more closely than SARIMA. This advantage stems from the model's time-varying transmission λ_t, which is explicitly modulated by paid marketing inputs and calendar effects (Figure 4), whereas SARIMA captures seasonality through an autoregressive structure alone.

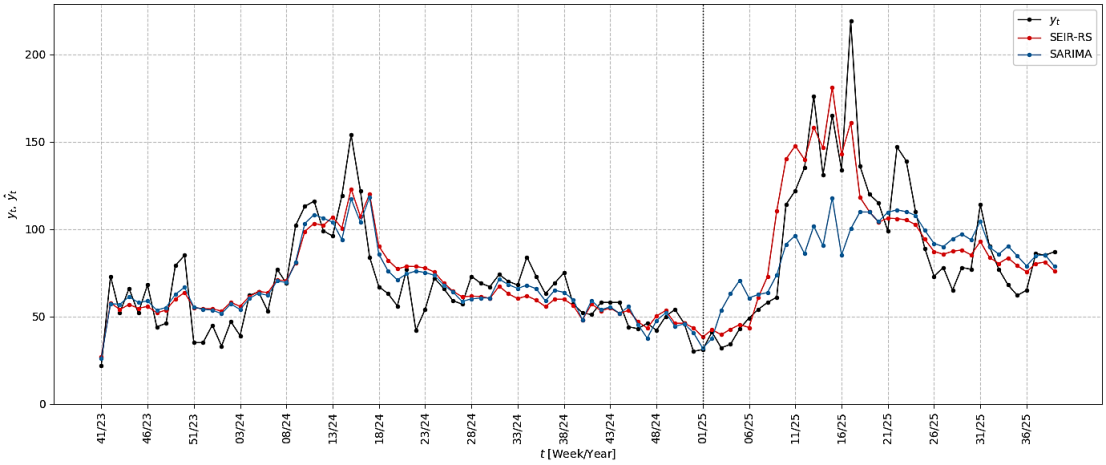

Figure 5. Observed weekly orders y_t (black) and forecasts from SEIRS (red) and SARIMA (blue). The vertical line marks the start of the test window (2025).

Table 2 summarises forecast errors on the test window. SEIRS improves all metrics relative to SARIMA: RMSE decreases by 14.4% (22.34 → 19.12), MAE by 15.9% (18.21 → 15.32), and MAPE by 33.3% (18.21% → 12.14%). The gains are most visible around local extrema (peaks and troughs), where SEIRS better captures changes in exposure and subsequent activation rather than extrapolating past levels. These results suggest that including latent life-cycle stages and varying transition strengths can yield more accurate short-term forecasts than a basic autoregressive seasonal model.

Table 2.

Comparison of forecast quality for SEIRS model and SARIMA in the test window (Figures 5, 6, $t \in (01/25,40/25)$; time step = 1 week)

Model	RMSE	MAE	MAPE
SEIRS	19.12	15.32	12.14%
SARIMA	22.34	18.21	18.21%

Figure 6 reports test-window residuals $e_t = y_t - \hat{y}_t$ for both models. SEIRS residuals cluster closer to zero with less apparent serial structure, while SARIMA exhibits larger errors around abrupt changes. The distributions are approximately symmetric, with occasional positive spikes consistent with short-lived shocks that neither method fully captures.

Figure 7 complements residual analysis with residual autocorrelation. For SEIRS (panel (a)), most $ACF(e_t)$ values lie within the 95% bounds, indicating limited remaining serial dependence after accounting for exposure–activation dynamics. For SARIMA (panel (b)), residual dependence persists beyond short lags, suggesting unmodelled structure related to latent behavioral transitions.

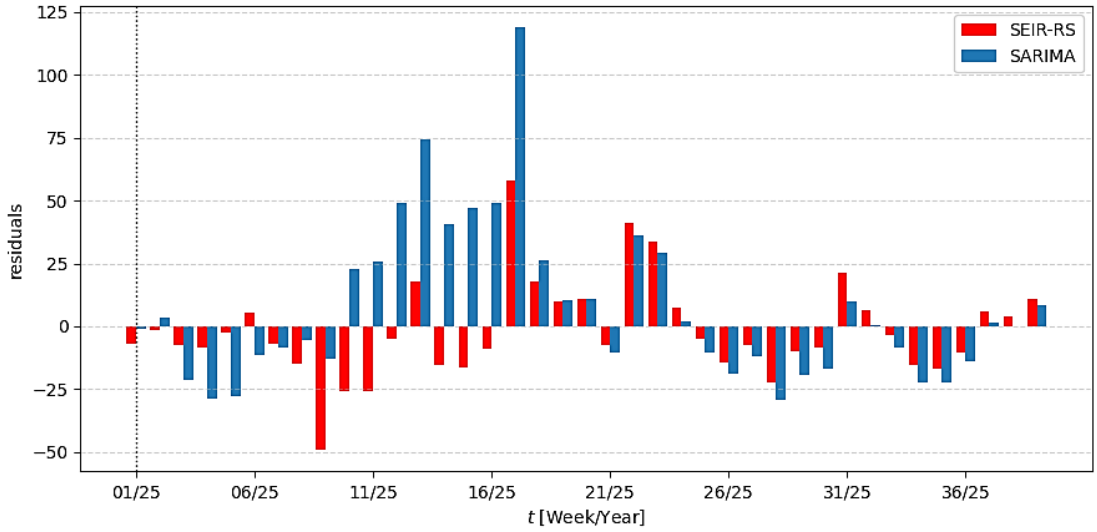

Figure 6. Test-window forecast residuals $e_t = y_t - \hat{y}_t$ for SEIRS (red) and SARIMA (blue).

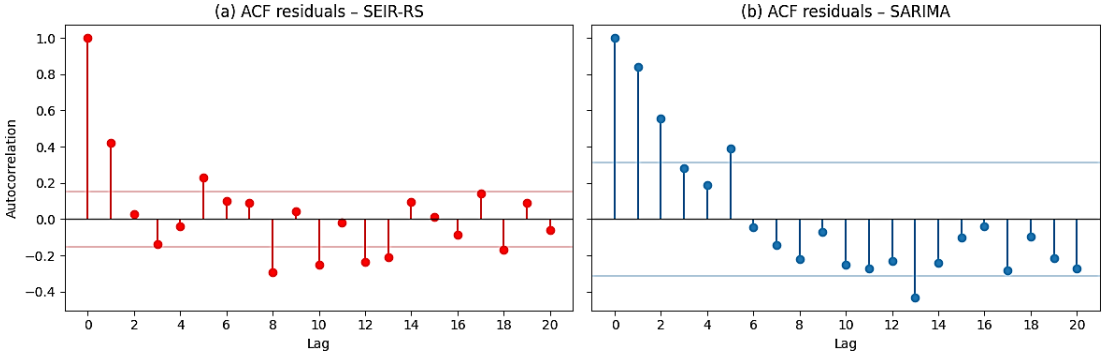

Figure 7. Residual autocorrelation $ACF(e_t)$ on the test window for (a) SEIRS and (b) SARIMA; the lighter lines indicate 95% confidence bounds.

4.2. SEIRS dynamics of the latent variables

Beyond forecasts, latent states provide a behavioral explanation of demand cycles. In particular, the trajectories of I_t (active) and R_t (temporarily inactive) individuals turn a black-box prediction into a life-cycle diagnosis by separating exposure effects from conversion and churn. The transmission factor λ_t acts as an early signal: it tends to rise before surges in activity and to recede after peaks, enabling diagnostic interpretation of interventions and seasonal shifts.

Figure 8 shows the latent compartments and transmission. The active state I_t co-moves with observed orders through the observation eq. (11). Importantly, increases in λ_t typically precede increases in I_t, consistent with a lead-lag mechanism in which shocks first affect exposure (via λ_t) and then translate into purchases through conversion σ. Together, the trajectories support interpreting demand peaks and lulls as the balance of transmission, churn γ, and reactivation ω under marketing and calendar modulation.

Figure 8. Inferred latent states and transmission for the SEIRS model: active buyers I_t, removed (inactive) state R_t as parts of N_0, and time-varying transmission λ_t over the training and test windows.

Figure 9 reports the demand momentum proxy $\mathfrak{R}_t = \lambda_t/\gamma$, with the horizontal reference line at $\mathfrak{R}_t = 1$. Periods with $\mathfrak{R}_t > 1$ indicate conditions under which exposure inflows are sufficient to offset churn, and they typically precede demand expansions. Conversely, sustained periods with $\mathfrak{R}_t < 1$ signal weakening demand and are associated with slowdowns unless marketing or seasonal factors lift λ_t. Interpreting $\mathfrak{R}_t$ jointly with Figure 8 helps distinguish peaks driven by propagation (high λ_t relative to γ) from peaks that require continued external lift.

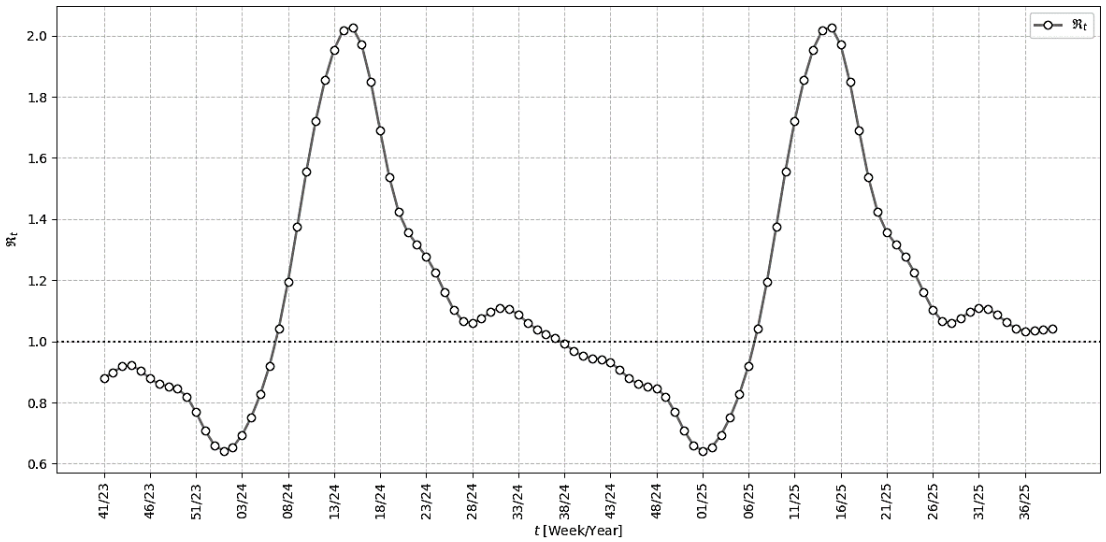

Figure 9. Demand-momentum proxy $\Re_t = \lambda_t/\gamma$ for the SEIRS model; the horizontal line indicates the threshold $\Re_t = 1$.

4.3. Sensitivity and scenario analysis

To assess robustness and planning relevance, we conduct a one-factor-at-a-time (OFAT) sensitivity analysis focused on short-horizon predictive accuracy. Each selected parameter is perturbed over a $\pm 20\%$ grid around its calibrated value while all remaining parameters are held fixed. For each setting, we report in-sample RMSE (RMSE train) and out-of-sample RMSE on the test window (RMSE final). Effect sizes are summarised with a tornado plot (Figure 10), and two scenario sweeps illustrate the qualitative demand implications for the most influential parameters (Figures 11-12).

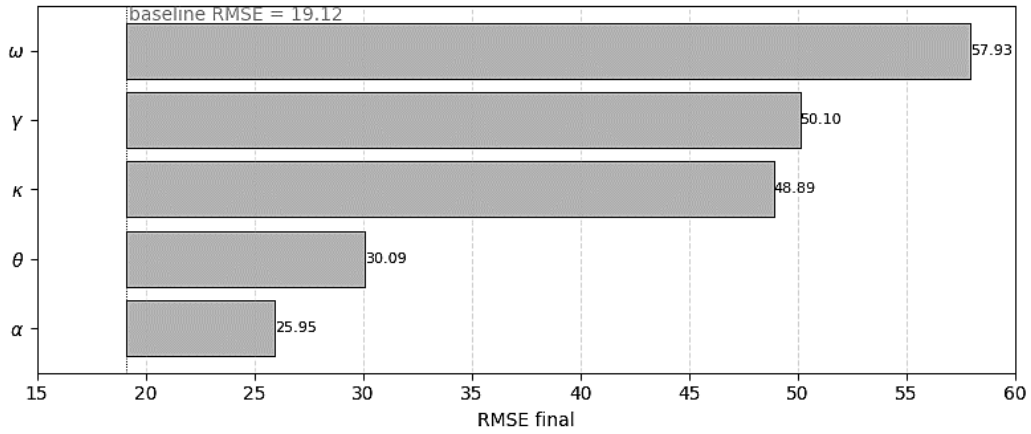

Figure 10. Tornado plot (one-way sensitivity) for RMSE final. Each bar shows the range of RMSE values from Table 3 after changing a single parameter, with the others remaining at their base level. The bars are sorted in descending order by width (the most significant impact at the top). The dashed line indicates the base value of RMSE final.

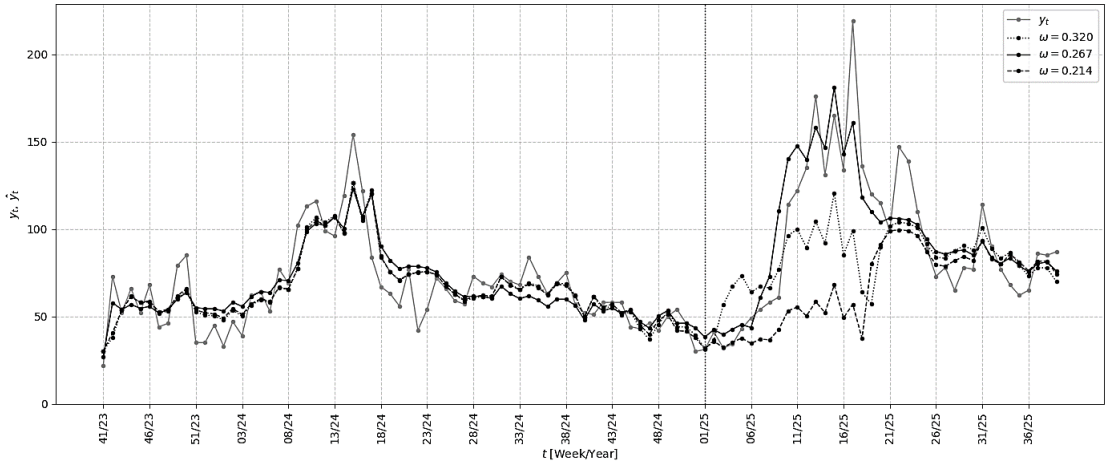

Figure 11. Forecast $\hat{y}_t$ for different values of reactivation factor: $\omega = 0.320, 0.267$ and 0.214 plotted by dotted, solid, and dashed lines, respectively. The observed values of weekly orders y_t is denoted by gray line.

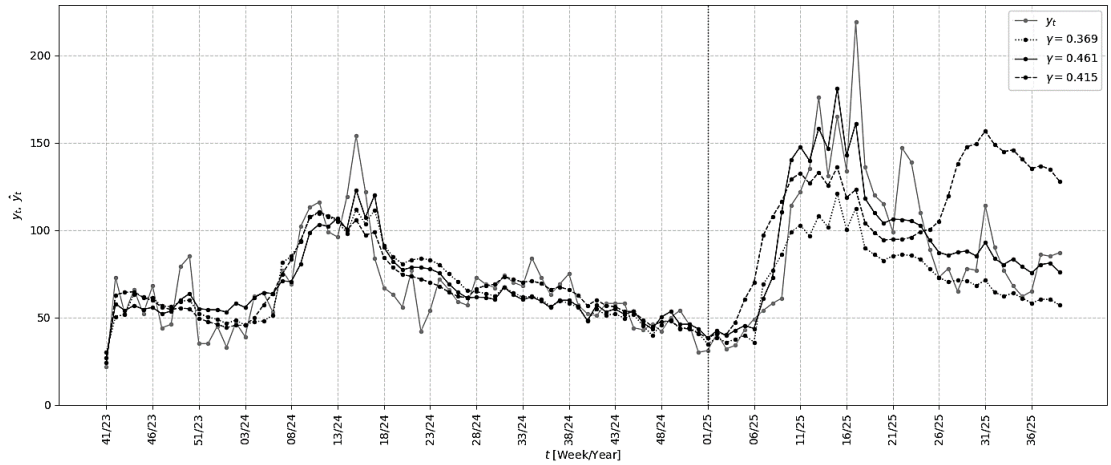

Figure 12. Forecast $\hat{y}_t$ for different values of churn factor: $\gamma = 0.369, 0.461$ and 0.415 plotted by dotted, solid, and dashed lines, respectively. The observed values of weekly orders y_t is denoted by gray line.

Table 3 and Figure 10 show that the test RMSE is most sensitive to the life-cycle transition rates: reactivation ω and churn γ. The scaling parameter κ is the next most influential parameter, reflecting that changes in per-active purchasing intensity translate directly into the observable y_t. More minor but still meaningful effects are associated with θ (seasonal/exogenous driver) and α (nonlinearity of social dependence), which primarily shape peak amplitude and curvature rather than changing the underlying momentum regime. The baseline test-window error is $RMSE_{\text{final}} = 19.1169$ (dashed reference line in Figure 10).

Table 3.
One-factor-at-a-time (OFAT) sensitivity of forecast accuracy. Each parameter is varied by $\pm 20\%$ *around its calibrated value (varied values are highlighted in gray) while others are fixed. The table reports train* $RMSE_{train}$ *and test windows* $RMSE_{test}$ *(higher values of* $RMSE_{test}$ *are highlighted in green).*

RMSE train	RMSE test	β_0	κ	ω	γ	δ	α	θ
1,519E+01	2,693E+01	-0,817	120	0,267	0,461	0,690	0,960	3,440
1,144E+01	2,505E+01	-0,817	110	0,267	0,461	0,690	0,960	3,440
1,379E+01	1,912E+01	-0,817	100	0,267	0,461	0,690	0,960	3,440
1,263E+01	2,410E+01	-0,817	90	0,267	0,461	0,690	0,960	3,440
1,284E+01	4,889E+01	-0,817	80	0,267	0,461	0,690	0,960	3,440
1,150E+01	5,793E+01	-0,817	100	0,320	0,461	0,690	0,960	3,440
1,258E+01	3,600E+01	-0,817	100	0,293	0,461	0,690	0,960	3,440
1,379E+01	1,912E+01	-0,817	100	0,267	0,461	0,690	0,960	3,440
1,278E+01	3,444E+01	-0,817	100	0,240	0,461	0,690	0,960	3,440
1,278E+01	3,901E+01	-0,817	100	0,214	0,461	0,690	0,960	3,440
1,257E+01	5,010E+01	-0,817	100	0,267	0,553	0,690	0,960	3,440
1,348E+01	4,853E+01	-0,817	100	0,267	0,507	0,690	0,960	3,440
1,379E+01	1,912E+01	-0,817	100	0,267	0,461	0,690	0,960	3,440
1,272E+01	2,529E+01	-0,817	100	0,267	0,415	0,690	0,960	3,440
1,258E+01	3,616E+01	-0,817	100	0,267	0,369	0,690	0,960	3,440
1,328E+01	2,263E+01	-0,817	100	0,267	0,461	0,690	1,150	3,440
1,402E+01	2,068E+01	-0,817	100	0,267	0,461	0,690	1,060	3,440
1,379E+01	1,912E+01	-0,817	100	0,267	0,461	0,690	0,960	3,440
1,331E+01	2,595E+01	-0,817	100	0,267	0,461	0,690	0,860	3,440
1,051E+01	2,565E+01	-0,817	100	0,267	0,461	0,690	0,780	3,440
1,269E+01	3,009E+01	-0,817	100	0,267	0,461	0,690	0,960	4,130
1,267E+01	2,492E+01	-0,817	100	0,267	0,461	0,690	0,960	3,780
1,379E+01	1,912E+01	-0,817	100	0,267	0,461	0,690	0,960	3,440
1,274E+01	2,287E+01	-0,817	100	0,267	0,461	0,690	0,960	3,100
1,274E+01	2,462E+01	-0,817	100	0,267	0,461	0,690	0,960	2,750

Figures 11 and 12 translate these relationships into demand (conversion) time history. Increasing reactivation ω accelerates replacement of the susceptible individuals pool ($R \to S$), leading to faster customer recoveries and higher predicted orders during volatile periods, whereas lower ω delays recovery and reduces demand rebuild. Varying churn γ primarily controls the persistence of activity: higher γ shortens active spells, truncates peaks, and produces faster post-peak decay; lower γ prolongs activity and raises plateau levels through stronger carry-over across weeks. Together, the OFAT ranking and scenario indicate that churn and win-back dynamics are first-order levers for both forecast performance and the qualitative shape of cycles, while θ and α mainly re-scale predictable seasonal amplitude and responsiveness to the active share.

5. Discussion, managerial implications, and conclusions

This case study shows that a state-space SEIRS e-commerce customer life-cycle model improves short-horizon demand forecasts and provides interpretable diagnostics for marketing- and seasonality-driven demand. In particular, the model outperforms a SARIMA benchmark and yields residuals with weaker serial dependence, suggesting that part of the temporal structure is captured via latent exposure–activation mechanisms rather than absorbed by autoregressive extrapolation. This result is consistent with the broader marketing diffusion tradition that emphasises structured mechanisms behind observed adoption or demand dynamics (Bass, 1969; Peres, Muller, Mahajan, 2010), and with evidence that epidemic-style thinking can be informative for diffusion processes under social influence (Fibich, 2016; McAdams, Song, 2025; Woo, Chen, 2016). In this regard, the empirical findings support **H1**.

The proposed latent states also support a behavioural interpretation of demand cycles that goes beyond purely statistical seasonality. Increases in the time-varying transmission intensity λ_t tend to precede increases in the active state I_t, consistent with a lead-lag mechanism: marketing pressure and seasonality first shift the exposed group and then translate into purchases through conversion. Similar lead-lag logic underpins surveillance-oriented state-space approaches in epidemiology, where latent intensities are tracked to anticipate subsequent changes in observed incidence (Papageorgiou, Tsaklidis, 2023; Prashad, 2025). The proxy of the demand growth rate (momentum) $\Re_t = \lambda_t/\gamma$ provides a compact diagnostic of whether exposure inflows are sufficient to offset churn weekly; sustained periods above (below) 1 correspond to expansive (recessive) configurations in terms of order numbers. This mirrors the role of the epidemic reproduction number as a regime indicator in epidemiology and its real-time estimation practice (Kermack, McKendrick, 1927; Nash, Nouvellet, Cori, 2022), thereby supporting **H2** and **H4** in the e-commerce management context.

Sensitivity analyses further clarify which mechanisms dominate in short-term matching and managerial leverage. Forecast error in the test window is most sensitive to life-cycle changes related to customer churn γ and reactivation ω, indicating that retaining and regaining customers determines the pace of recovery after peaks and slowdowns, while other parameters primarily reshape predictable seasonal amplitude and responsiveness to conversion. This complements applied churn research, where predictive pipelines are often framed around identifying customers at risk of leaving (Dias, Antonio, 2025), and aligns with recent emphasis on explainability when churn-related outputs are used for intervention timing and decision support (Kimitei et al., 2025). Unlike churn treated as a standalone classifier, the SEIRS formulation integrates customer loss and reactivation directly into the demand system, consistent with the hypothesis that these mechanisms influence demand trajectories, supporting **H3**.

This case study shows that a state-space SEIRS e-commerce customer life-cycle model improves short-horizon demand forecasts and provides interpretable diagnostics for marketing- and seasonality-driven e-commerce demand. Empirically, the forecast outperforms a SARIMA benchmark and leaves residuals with weaker serial dependence, consistent with capturing exposure–activation dynamics rather than extrapolating past levels.

6. Conclusions

The proposed state-space SEIRS e-commerce framework is a marketing analysis tool that combines weekly orders with an e-commerce customer life-cycle model. It segments customers by life-cycle stage, improving trend prediction and strategic planning for customer retention.

Two implications are beneficial. In marketing campaign management, stabilizing the number of orders is mainly a retention and reactivation task: lowering churn (γ) lengthens active spells and slows post-peak orders number decay, while higher reactivation (ω) speeds recovery by replenishing the susceptible pool. Paid media works best when λ_t and $\mathfrak{R}_t$ increase, higher spend is more likely to be amplified, whereas when $\mathfrak{R}_t < 1$, short external impulses may be necessary to restore dynamics.

Operationally, a compact dashboard can track time-varying transmission intensity λ_t as an early-warning signal, the value of the demand-momentum proxy $\mathfrak{R}_t$ (relative to 1), and churn/reactivation (γ, ω). Short-term forecasts and residual checks should appear alongside to signal emerging deviations from expected values and guide paid media spending. In the weekly planning–execution–review cycle, λ_t and R_t set the pace, while γ and ω set retention and recovery priorities, creating a shared view of demand shifts. Following the distinction between real-time monitoring and predictive analytics in organisational operations (Wolniak, 2023), the proposed indicators can be positioned as a compact monitoring layer that complements forecasting.

This study has several limitations. It relies on weekly aggregates from a single retailer and uses a limited representation of marketing exposure and other exogenous drivers. The model also assumes a homogeneous population and captures channel effects in reduced form, which limits campaign-level causal inference. Future research will validate the approach on additional cases (other retailers and product categories) to assess external validity and boundary conditions. Further extensions include richer driver constructions (e.g., adstock/carry-over, multi-channel structure, moving-holiday effects) and probabilistic uncertainty reporting for latent states and indicators.

Declarations

Funding: This research received no external funding.

Data availability: The dataset is proprietary retailer data and cannot be publicly shared. Aggregated series or derived indicators may be provided upon reasonable request. Conflict of interest: The authors declare no conflict of interest.

References

1. Bass, F.M. (1969). A New Product Growth for Model Consumer Durables. *Management Science*, *15*(5), 215-227. https://doi.org/10.1287/mnsc.15.5.215
2. Dias, J.R., Antonio, N. (2025). Predicting customer churn using machine learning: A case study in the software industry. *Journal of Marketing Analytics*, *13*(1), 111-127. https://doi.org/10.1057/s41270-023-00269-9
3. Fibich, G. (2016). Bass-SIR model for diffusion of new products in social networks. *Physical Review E*, *94*(3), 032305. https://doi.org/10.1103/PhysRevE.94.032305
4. Kermack, W.O., McKendrick, A.G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *115*(772), 700-721. https://doi.org/10.1098/rspa.1927.0118
5. Kimitei, S., Agiro, D., Ni, S., Ni, H. (2025). Predictability & explainability of survival analysis in churn prediction. *Journal of Marketing Analytics*. https://doi.org/10.1057/s41270-025-00450-2
6. McAdams, D., Song, Y. (2025). Adoption epidemics and viral marketing. *Theoretical Economics*, *20*(2), 453-480. https://doi.org/10.3982/TE5886
7. Medrek, M., Pastuszak, Z. (2021). Numerical simulation of the novel coronavirus spreading. *Expert Systems with Applications*, *166*. https://doi.org/10.1016/j.eswa.2020.114109
8. Nash, R.K., Nouvellet, P., Cori, A. (2022). Real-time estimation of the epidemic reproduction number: Scoping review of the applications and challenges. *PLOS Digital Health*, *1*(6), e0000052. https://doi.org/10.1371/journal.pdig.0000052
9. Papageorgiou, V.E., Tsaklidis, G. (2023). An improved epidemiological-unscented Kalman filter (hybrid SEIHCRDV-UKF) model for the prediction of COVID-19. Application on real-time data. *Chaos, Solitons, and Fractals*, *166*, 112914. https://doi.org/10.1016/j.chaos.2022.112914

10. Peres, R., Muller, E., Mahajan, V. (2010). Innovation diffusion and new product growth models: A critical review and research directions. *International Journal of Research in Marketing*, *27*(2), 91-106. https://doi.org/10.1016/j.ijresmar.2009.12.012
11. Prashad, C.D. (2025). State-space modelling for infectious disease surveillance data: Stochastic simulation techniques and structural change detection. *Infectious Disease Modelling*, *10*(4), 1507-1532. https://doi.org/10.1016/j.idm.2025.05.005
12. Rogers, E. (2003). *Diffusion of innovations : 5th ed*. Free Press.
13. Tillett, H.E. (1992). Infectious Diseases of Humans: Dynamics and Control. In: R.M. Anderson, R.M. May (eds.), *Epidemiology and Infection*, *108*(1) (pp. 211-211). Oxford University Press; 1991. https://doi.org/10.1017/S0950268800059896
14. Wolniak, R. (2023). Functioning of real-time analytics in business. *Scientific Papers of Silesian University of Technology Organization and Management Series*, *172*. https://doi.org/10.29119/1641-3466.2023.172.40
15. Wolniak, R. (2024). Analyzing customer behavior – employing business analytics within Industry 4.0 ecosystems. *Scientific Papers of Silesian University of Technology Organization and Management Series*, *195*. https://doi.org/10.29119/1641-3466.2024.195.39
16. Woo, J., Chen, H. (2016). Epidemic model for information diffusion in web forums: experiments in marketing exchange and political dialog. *SpringerPlus*, *5*(1), 66. https://doi.org/10.1186/s40064-016-1675-x