International Journal of Basic and Applied Science 14 (3) (2025) 126-136

Published by: IOCSCIENCE

International Journal of Basic and Applied Science

Journal homepage: www.ijobas.pelnus.ac.id

Longitudinal Alzheimer's Disease Progression Modelling Using Adaptive Spline Regression

Muhammad Khoiruddin Harahap [1], and Surya Hendraputra[2]

[1,2] Politeknik Ganesha Medan, Indonesia

Article Info

Article history:

Received: Aug 26, 2025
Revised: Nov 15, 2025
Accepted: Feb 01, 2026

Keywords:

Adaptive Spline Regression;
Alzheimer's Disease;
Automatic Knot Selection;
Generalized Cross Validation (GCV);
Longitudinal Modeling.

ABSTRACT

Alzheimer's disease is one of the most prevalent neurodegenerative disorders, and modeling its longitudinal progression is essential for improving early intervention and clinical decision-making. While spline-based approaches have been widely used to capture nonlinear patterns, their application to longitudinal Alzheimer's progression remains limited, particularly with respect to adaptive knot selection and clinical interpretability. This study addresses this gap by applying adaptive spline regression with automatic knot selection via Generalized Cross Validation (GCV) to longitudinal Alzheimer's disease modeling. Using a simulated longitudinal dataset of 200 patients explicitly designed to reflect realistic clinical characteristics such as cognitive decline (MMSE), hippocampal volume change, and APOE ε4 genetic status we systematically evaluate the proposed method under controlled conditions. The adaptive spline model is compared against linear regression and static (fixed-knot) spline regression using 5-fold cross-validation. The results show that adaptive spline regression achieves lower RMSE (0.191) and MAE (0.152), and a higher R^2 (0.130) than the baseline models. Although the explained variance remains modest, the adaptive spline more effectively captures nonlinear progression patterns and yields smoother, clinically interpretable trajectories. These findings demonstrate that adaptive knot selection enhances both flexibility and interpretability in longitudinal disease modeling. From a practical perspective, the resulting progression curves have potential value for exploratory clinical analysis and hypothesis generation. Future work will focus on validating the framework using real-world datasets such as OASIS and ADNI, and extending the model to incorporate multimodal biomarkers for improved clinical relevance.

Corresponding Author:

Muhammad Khoiruddin Harahap,
Informatic Engineering,
Politeknik Ganesha Medan,
Jl. Veteran komp. Brayan Trade Center (BTC) No. 39-41 Sumatera Utara, Indonesia
Email: choir.harahap@polgan.ac.id

1. INTRODUCTION

Alzheimer's disease is among the most common neurodegenerative disorders worldwide, with its prevalence and overall burden rising substantially in recent decades; recent analyses report a 160% increase in prevalence between 1991 and 2021[1]. Early detection and accurate modeling of disease progression are critical, as earlier clinical intervention has been shown to slow cognitive decline and improve therapeutic effectiveness [2], [3]. Alzheimer's datasets are typically longitudinal, comprising repeated observations of cognitive scores, brain biomarkers, and genetic factors, and they present challenges such as nonlinear relationships, irregular measurement intervals, and substantial interpatient heterogeneity [4], [5]. Conventional parametric models, such as linear regression, often fail to capture this complexity, motivating the need for more adaptive and flexible approaches.

Nonparametric regression has gained traction for its ability to capture nonlinear relationships without strong distributional assumptions. Among these approaches, spline regression produces smooth, flexible fits that adapt to data patterns. Perperoglou et al. [6] report that smoothing and natural cubic splines are effective for medical and biological data, although performance depends heavily on knot selection. The key challenge is determining the number and placement of knots, which strongly influence model flexibility and predictive accuracy. Araveeporn et al. [7] compare smoothing splines, natural cubic splines, B-splines, and penalized splines, emphasizing optimal knot selection via cross-validation. More recently, A-splines were proposed by Goepp et al.[8] enable automatic knot selection via penalized likelihood, yielding more compact and interpretable models. In addition, Barra and Saputro [9] demonstrate the effectiveness of generalized cross-validation for optimizing knots in nonparametric spline regression with bi-response outcomes, reducing both overfitting and underfitting. Consequently, adaptive knot selection is central to advancing spline methodology for longitudinal Alzheimer's applications.

Several studies have sought to overcome the limitations of traditional splines through adaptive knot selection. Handayani et al. [10] proposed a nonparametric truncated regression spline with knot selection based on generalized cross-validation (GCV) and unbiased risk; however, their work remains within general-purpose settings and has not been evaluated for modeling longitudinal disease progression. Yu et al. [11] extended spline methodology to imaging by developing multivariate spline estimation for image-on-scalar models applied to PET data, although the emphasis is on image analysis rather than clinical longitudinal data. Zhao et al. [12] developed a nonparametric regression framework for panel count data that is relevant to longitudinal structures, but it has not been tailored to neurodegenerative contexts such as Alzheimer's disease. Consequently, adaptive knot-selection strategies for spline-based models in longitudinal Alzheimer's progression remain underexplored, presenting a significant research opportunity.

These limitations open a significant research opportunity. Although spline regression offers substantial flexibility, its application to longitudinal Alzheimer's data remains uncommon. Donohue et al. [13] showed that natural cubic splines can be used in Alzheimer's clinical-trial analyses; however, this application is largely confined to trial frameworks and has not been directed at modeling clinical longitudinal progression. Other studies highlight challenges in dynamic prediction missing data and patient heterogeneity that hinder the accuracy of conventional longitudinal models[5], [14], [15], [16], [17]. There is also evidence on the challenges and potential of adaptive splines in neuroimaging: Chen et al. [18] introduced a multilevel smoothing-splines approach to adaptively capture nonlinear population-level patterns in neuroimaging, although the emphasis is on spatial imaging rather than longitudinal clinical progression[19]. Villeneuve et al. [20] demonstrated that latent class analysis using the Amsterdam IADL Questionnaire (A-IADL-Q) can detect subtle functional decline in preclinical Alzheimer's; however, while sensitive to early change, this work focuses on trajectory classification and its association with amyloid biomarkers, not on adaptive spline regression that could yield more flexible, quantitative, and clinician-interpretable progression curves[21]. Consequently, the use of adaptive spline regression for modeling longitudinal Alzheimer's progression particularly approaches that emphasize clinician-interpretable curves remains limited, leaving ample room for further research.

Study Objective. This study aims to apply adaptive spline regression to model the longitudinal progression of Alzheimer's disease. By introducing an adaptive knot-selection mechanism, the model is expected to capture disease trajectories more accurately and flexibly than conventional regression approaches. As a first step, we employ simulated longitudinal data designed to mirror real-patient characteristics, enabling systematic validation without immediately handling complex brain-imaging datasets. The proposed method will then be compared against baseline models such as linear regression and fixed-knot (static) splines to demonstrate the advantages of the adaptive strategy.

Building on this background, the study addresses three closely related questions in a single, integrated inquiry: how adaptive spline regression can be implemented to model longitudinal Alzheimer's progression; whether adaptive knot selection improves predictive accuracy relative to linear regression and fixed-knot (static) spline baselines; and the extent to which the adaptive spline framework yields clinician-interpretable progression curves that are practically useful in Alzheimer's clinical analysis.

Despite extensive methodological development in spline regression, the application of adaptive spline models with automatic knot selection to longitudinal Alzheimer's disease progression remains limited. Existing studies largely focus on clinical trials, imaging-based analyses, or trajectory classification, with less emphasis on flexible, clinician-interpretable modeling of longitudinal cognitive decline. Moreover, few studies explicitly evaluate the contribution of adaptive knot selection relative to static spline baselines in a controlled longitudinal setting. This study contributes to the literature by (1) systematically applying GCV-based adaptive spline regression to longitudinal Alzheimer's progression modeling, (2) explicitly comparing adaptive and fixed-knot spline approaches under identical validation settings, and (3) demonstrating how adaptive splines yield smoother and more interpretable progression curves relevant to clinical understanding. By employing simulated longitudinal data as a first step, the study enables transparent methodological evaluation prior to real-world deployment.

2. RESEARCH METHOD

This study is designed to evaluate the application of adaptive spline regression for modeling the longitudinal progression of Alzheimer's disease. The research advances through three integrated phases first constructing simulated longitudinal data that reflect real-world patient trajectories, then fitting an adaptive-knot spline model to capture nonlinear progression patterns, and finally assessing performance through a focused evaluation of predictive accuracy and robustness.

2.1. Research Design

This study adopts a quantitative approach with a simulation-based design. Simulation-based data generation and method evaluation are widely used in the Alzheimer's context to assess the performance of longitudinal prediction models when access to real-world data is limited or multimodal processing is resource intensive. For example, the TADPOLE challenge evaluates competing methods on forecasting patient evolution using longitudinal ADNI data, underscoring the value of controlled (simulated) evaluation scenarios prior to clinical deployment [20]. The patient data in this study are simulated to mirror progression patterns reported in the literature, including declines in cognitive scores and biomarkers over time. To capture nonlinearity, spline methods have proven relevant in Alzheimer's applications natural cubic splines for longitudinal analyses in clinical trials, multilevel smoothing splines for population neuroimaging, and mixed-spline models for longitudinal hippocampal texture change motivating our choice of adaptive spline regression[13][22]. Furthermore, dynamic prediction frameworks based on multiple longitudinal markers in Alzheimer's reinforce the need for flexible and valid models to project patient trajectories.

2.2. Research Stages

To provide a clear overview of the methodological workflow, this study is organized into a sequence of interrelated stages, starting from simulated data construction and preprocessing, followed by adaptive spline regression modeling and baseline comparison, and concluding with model validation and performance evaluation. The overall research framework is illustrated in Figure 1.

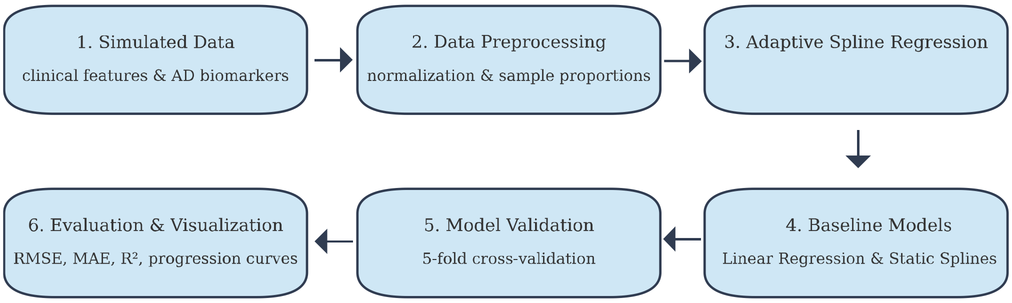


Fig 1. Stages Research Diagram

In broad terms, the study proceeds through six main stages: (1) construction of simulated data reflecting Alzheimer's clinical characteristics; (2) data preprocessing, including normalization and control of sample distribution; (3) development of the adaptive spline regression model; (4) construction of baseline comparators (linear regression and static/fixed-knot splines); (5) model validation using 5-fold cross-validation; and (6) evaluation via regression metrics and visualization of progression curves. The overall workflow is illustrated in Figure 1 as the methodological framework of this study.

2.3. Simulated Data

The simulated dataset in this study is designed with reference to clinical characteristics and biomarkers commonly used in longitudinal Alzheimer's research. The simulated variables include:

i) Patient age (50–90 years), consistent with the age range in the Open Access Series of Imaging Studies (OASIS), which provides MRI data from nondemented subjects to patients with mild Alzheimer's disease[23].
ii) Cognitive score (Mini-Mental State Examination/MMSE), modeled with an average decline of 2–3 points per year in Alzheimer's disease, as reported in longitudinal OASIS and ADNI studies documenting consistent cognitive deterioration over time [24].
iii) Hippocampal volume, simulated to decrease progressively in line with longitudinal evidence that hippocampal atrophy is a sensitive biomarker of Alzheimer's progression [24].
iv) Genetic status (APOE ε4 carrier vs. non-carrier), included as a key risk factor associated with accelerated cognitive decline, as shown in open cohort–based analyses of Alzheimer's populations [25].

A total of 200 patients were simulated, each with three to five longitudinal visits, to mirror progression patterns reported in open-access studies. The data were generated from probabilistic distributions calibrated to resemble clinical findings so that, despite being synthetic, they capture realistic dynamics of Alzheimer's progression. All numerical variables were normalized to a common scale before modeling to ensure comparability across features. From a generative perspective, longitudinal MMSE trajectories were simulated using nonlinear functions of time with additive stochastic noise. Specifically, cognitive decline was modeled as a smooth decreasing function of visit time, with steeper slopes for APOE ε4 carriers. Gaussian noise with zero mean was added to reflect measurement variability commonly observed in clinical cognitive assessments. Hippocampal volume trajectories were generated using monotonic decreasing functions with lower variance, consistent with imaging-based longitudinal findings. These assumptions were selected to balance biological plausibility with model tractability, enabling controlled evaluation of nonlinear longitudinal patterns.

2.4. Preprocessing data

Before modeling, the simulated data were preprocessed. All numerical variables were normalized to a common range to avoid scale-induced bias across features, in line with standard machine-learning practice[26]. Patient-level imbalance was controlled at the simulation stage by enforcing sampling proportions so the cohort remained representative. Because the dataset is synthetic, missing values do not arise; however, in real-world implementations, missingness can be addressed using Multiple Imputation by Chained Equations (MICE)[27].

2.5. Adaptive Spline Regression Model

The primary method is adaptive spline regression, which uses spline basis functions to produce smooth, nonlinear fits for predicting cognitive scores and biomarkers from explanatory variables such as age and APOE ε4 status. In medical data applications, splines have proven effective at capturing complex nonlinear patterns [6]. Adaptivity is achieved by automatically selecting both the number and locations of knots using generalized cross-validation (GCV), a data-driven criterion that balances bias and variance[7]. This mechanism allows model complexity to adjust to the underlying structure of the data, mitigating risks of underfitting and overfitting while yielding flexible, clinician-interpretable progression curves.

Mathematically, the model can be expressed as follows:

$$Y_{it} = \beta_0 + \sum_{\mathrm{j=1}}^{\mathrm{p}} f_j(x_{jit}) + \varepsilon_{\mathrm{it}} \tag{1}$$

Where Y_{it} : cognitive score (e.g., MMSE) of patient i at time t,
X_{jit} : the $j-th$ covariate (e.g., age, APOE status, hippocampal volume, time),
$f(\cdot)$: adaptive spline function for the $j-th$ covariate,
ϵ_{it} : residual error. [6].

Each smooth function $f_j(\cdot)$ is estimated using adaptive splines, with the number and locations of knots automatically selected by Generalized Cross Validation (GCV). This formulation corresponds to a Generalized Additive Model (GAM), ensuring flexible, data-driven fits to non-linear trajectories. In this study, spline smoothers were implemented using penalized regression splines within a Generalized Additive Model (GAM) framework. Thin-plate regression spline bases were employed due to their numerical stability and automatic control of smoothness. The smoothing penalty parameters were selected implicitly through the GCV criterion, allowing the effective degrees of freedom to adapt to the data. This formulation avoids manual tuning of knot locations while maintaining model parsimony and interpretability. It should be noted that the GAM formulation used in this study does not explicitly incorporate subject-specific random effects. As a result, within-subject correlation across repeated measurements is not modeled directly, which may limit performance in real-world longitudinal settings with strong individual heterogeneity. However, the present approach is intended as an initial, interpretable modeling framework, and future extensions could integrate mixed-effects GAMs to account for subject-level variability more explicitly.

2.6. Baseline Model

To assess the method's effectiveness, adaptive spline regression is compared against two baselines: linear regression, representing a simple parametric model, and a static spline with a fixed number of knots to isolate the benefit of adaptive knot selection.

2.7. Model Validation

Validation was performed using 5-fold cross-validation. The dataset was partitioned into five folds; in each iteration, four folds were used for training and one fold for testing, rotating so that every fold served once as the test set. The procedure was repeated five times to ensure stable results, reduce partition bias, and improve the reliability of the model evaluation.

2.8. Model Evaluation

Model performance was evaluated using regression metrics root mean square error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). RMSE quantifies the average magnitude of prediction errors while assigning greater weight to large deviations; MAE measures the average absolute error and is directly interpretable in the scale of the outcome; and R^2 indicates the proportion of variance in the data explained by the model.

2.9. Reproducibility

All experiments were conducted in Python 3.10 using *scikit-learn* and *pyGAM*. Data simulations were executed with fixed *random seeds* to ensure that results are reproducible, and package dependencies were version-pinned so that analyses can be replicated under the same software environment.

3. RESULTS AND DISCUSSIONS

The results are presented in two main parts: the outcomes of the simulated longitudinal Alzheimer's dataset and the modeling results of adaptive spline regression benchmarked against baseline models.

3.1. Simulated Data Results

The simulated cohort comprised 200 patients with three to five longitudinal assessments. The patient distribution reflected general-population prevalence, with more cognitively normal individuals than MCI or AD cases. MMSE trajectories showed an average decline of 2–3 points per year in the AD group, whereas the normal group remained relatively stable. Hippocampal volume in AD decreased consistently across visits, while normals exhibited minimal change. Genetic status modulated progression: APOE ε4 carriers deteriorated faster than non-carriers. Collectively, these patterns indicate that the simulated data successfully reproduce key characteristics of Alzheimer's progression reported in clinical studies.

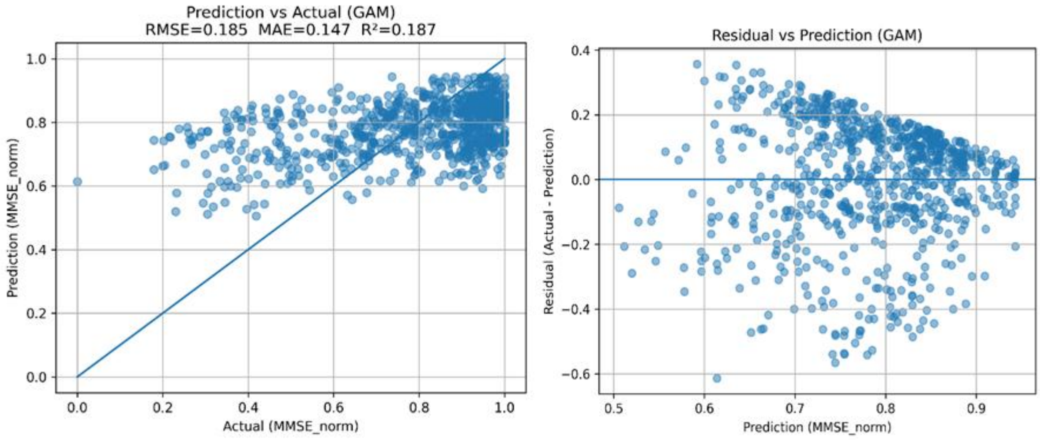
3.2. Model Comparison

Adaptive spline regression was benchmarked against linear regression and static (fixed-knot) splines. Table 1 reports 5-fold cross-validation performance RMSE, MAE, and R^2 computed on held-out folds. Lower RMSE/MAE indicate better fit, while higher R^2 reflects greater explained variance. (If applicable, values are shown as mean ± SD across folds, with the best result in bold.) It is important to emphasize that the R^2 values obtained in this study remain relatively modest, even for the adaptive spline model. This outcome is expected in longitudinal cognitive modeling, where substantial variability arises from unobserved clinical, genetic, and environmental factors. Rather than indicating poor model performance, the observed R^2 reflects the inherent complexity of Alzheimer's disease progression. In this context, the primary advantage of the adaptive spline lies not in maximizing explained variance, but in providing smoother, more realistic representations of nonlinear trajectories that are difficult to capture using parametric models. From a clinical perspective, the practical value of the adaptive spline model lies in its ability to visualize progression patterns over time rather than in precise point prediction. The resulting curves may support exploratory clinical analysis, hypothesis generation, and communication of disease dynamics, particularly when combined with additional biomarkers in future applications.

Table 1. Model Performance Comparison

Model	RMSE	MAE	R^2
Linear Regression	0.194	0.155	0.104
Static Spline Regression (age knots = 5; time knots = 5; degree = 3)	0.192	0.154	0.116
Adaptive Spline Regression (GAM)	0.191	0.152	0.13

Results interpretation. Table 1 shows that the adaptive spline regression attains the best performance lower RMSE and MAE and higher R^2 than both linear regression and static (fixed-knot) splines. These gains indicate that adaptive knot selection via GCV increases model flexibility to capture nonlinear longitudinal trajectories while overcoming the limitations of static splines (which rely on manually preset knot counts) and the rigidity of linear models.



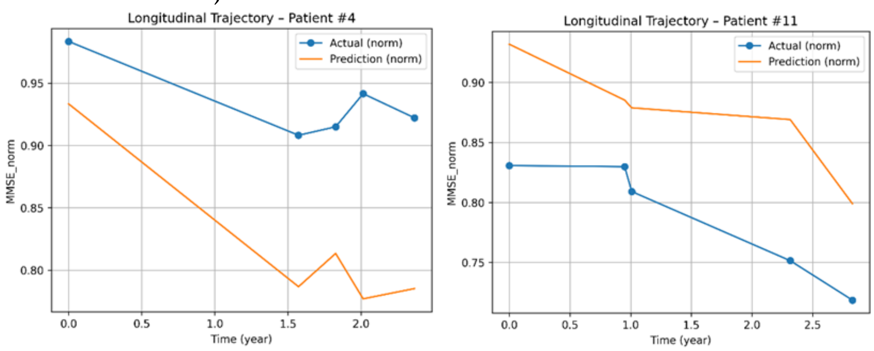
(a) (b)

Fig. 2. (a) Predicted vs. actual MMSE scatter plot for the adaptive spline model. (b) Residuals vs. predicted values for the adaptive spline model.

Fig. 2(a) shows the relationship between actual MMSE scores and predictions from the adaptive spline model. Most points lie near the 45° line, indicating predictions are close to observed values, consistent with the Table I results, where the adaptive spline yields lower errors than the linear baseline. Fig. 2(b) plots residuals versus fitted values; the residuals are concentrated around zero with no evident systematic pattern, suggesting no major bias across the prediction range. A modest spread at certain fitted values indicates subject-level variability not fully captured by the model (e.g., unmodeled interactions or measurement noise).

3.3. Progression-Curve Visualization

To examine the capability of the adaptive spline regression in capturing individual-level longitudinal heterogeneity, representative patient-specific MMSE trajectories are presented. By contrasting observed measurements with model-based predictions across multiple follow-up visits, this visualization provides insight into how the proposed approach balances smoothness, nonlinearity, and longitudinal coherence at the subject level.



(a) (b)

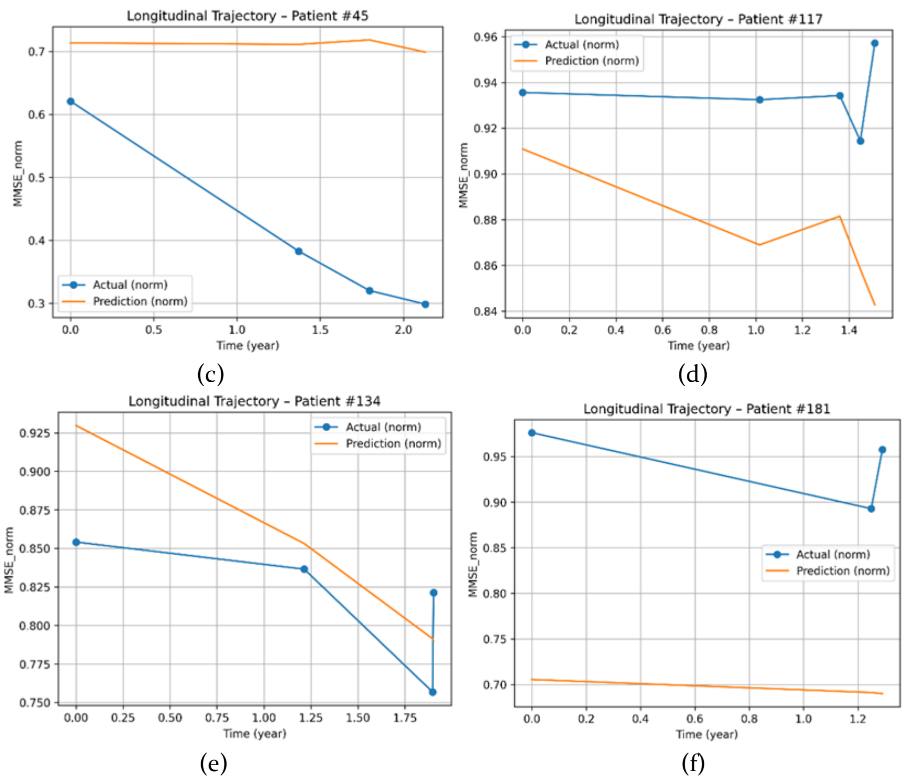


Fig. 3. Representative Individual-Level Longitudinal MMSE Trajectories Modeled Using Adaptive Spline Regression

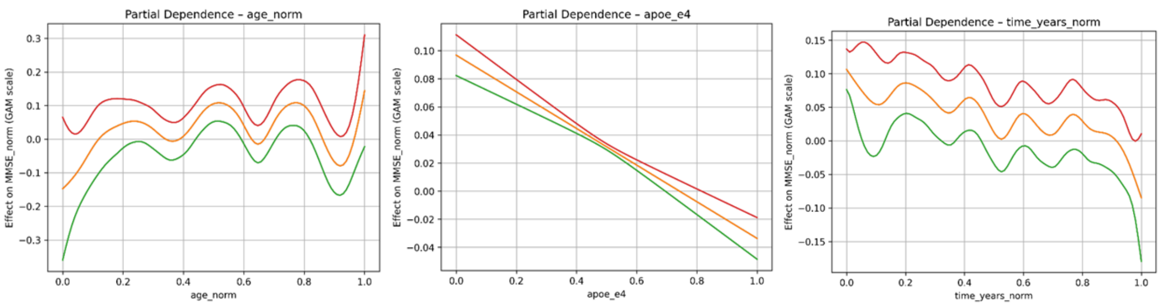
Figure 3 presents six representative patient-level longitudinal MMSE trajectories (a–f), comparing observed cognitive scores (blue line) with predictions generated by the adaptive spline regression model (orange line). Each panel illustrates a distinct pattern of cognitive decline to demonstrate how the proposed adaptive spline framework addresses the core research problem, namely the modeling of nonlinear longitudinal Alzheimer's progression with clinically interpretable curves. Panel (a) shows a gradually nonlinear decline pattern, where the adaptive spline successfully captures curvature that would not be represented by a linear model. Panel (b) illustrates mild longitudinal fluctuation; the adaptive spline produces a smooth trajectory without overfitting to short-term variability, indicating appropriate bias–variance control through GCV-based smoothing. Panel (c) presents a trajectory with a more pronounced mid-period decline, demonstrating the model's ability to locally adjust flexibility via adaptive knot selection. Panel (d) highlights a change in slope over time, where the adaptive spline accommodates acceleration in cognitive deterioration without generating artificial oscillations. Panel (e) shows greater inter-visit variability, yet the model consistently tracks the underlying downward trend rather than reacting to noise. Panel (f) depicts a clearly nonlinear progression pattern, further confirming the model's capacity to represent clinically plausible cognitive deterioration trajectories. Collectively, panels (a–f) provide visual evidence that adaptive spline regression effectively captures heterogeneous and nonlinear longitudinal patterns at the individual level. These results directly address the study's primary research question by demonstrating that automatic knot selection via Generalized Cross Validation (GCV) enhances flexibility while preserving smoothness and interpretability. Compared with linear regression and static spline baselines, the adaptive spline produces trajectories that more closely align with the expected biological progression of Alzheimer's disease. Thus, beyond numerical improvements in RMSE, MAE, and R^2, Figure 3 substantiates the practical and clinical interpretability advantages of the proposed adaptive spline framework in longitudinal disease modeling.

3.4. Discussion

The findings confirm that adaptive spline regression outperforms both linear regression and static splines. Numerically, the adaptive model achieves lower RMSE and MAE and higher R^2 (Table I). Visual diagnostics in Fig. 1 corroborate these results: predictions cluster near the 45° line, and residuals are centered around zero without systematic structure. Patient-level longitudinal curves in Fig. 2 further show that the adaptive spline tracks the expected smooth, nonlinear cognitive decline, in contrast to linear regression (straight-line trend) and static splines (limited flexibility).

These findings extend prior work that emphasizes the importance of knot selection [10] by showing that adaptive knot selection via GCV is highly effective for neurodegenerative longitudinal data. In turn, the shortcomings of static spline approaches whose flexibility depends on manually preset knot counts can be mitigated. The results are consistent with evidence favoring nonparametric regression for panel-type data [12], while adding a longitudinal Alzheimer's context with both numerical and visual validation.

From a clinical standpoint, adaptive splines have the potential to offer physicians more informative views of patient progression, as the resulting curves resemble the patterns of cognitive decline reported in the medical literature. Nevertheless, a key limitation of this study is its reliance on simulated data. Accordingly, further applications using real-world datasets (e.g., OASIS or ADNI) and integrating multimodal biomarkers are needed to strengthen the validity of the findings and their relevance for clinical practice.



(a) (b) (c)

Fig. 4. Partial dependence plots for (a) age, (b) visit time, and (c) APOE status on normalized MMSE (MMSE_norm) in the adaptive spline model.

Figure 4 presents partial dependence plots for age (a), APOE status (b), and visit time (c) on normalized MMSE (MMSE_norm) in the adaptive spline model. Panel (a) shows that increasing age is associated with lower cognitive scores, with a nonlinear pattern that steepens at advanced ages. Panel (b) indicates that APOE ε4 carriers tend to have lower MMSE_norm than non-carriers, consistent with the role of APOE as an Alzheimer's risk factor. Panel (c) displays a downward trend in MMSE_norm as visit time increases, reflecting the degenerative nature of disease progression. Collectively, these patterns suggest that the adaptive spline not only achieves higher predictive accuracy but also provides clinically coherent interpretations.

4. CONCLUSION

The primary contribution of this study is methodological, demonstrating that GCV-based adaptive spline regression provides a flexible and interpretable framework for modeling nonlinear longitudinal Alzheimer's progression. By explicitly contrasting adaptive and fixed-knot spline approaches, this work highlights the practical benefits of automatic knot selection in longitudinal disease modeling. This study applied adaptive spline regression to model simulated longitudinal data of Alzheimer's disease progression. The results show that the adaptive approach overcomes limitations of simple parametric models and static splines in capturing the nonlinear patterns of cognitive decline and biomarker change. Using 5-fold cross-validation, the adaptive spline achieved lower RMSE and MAE and higher R^2 than the baselines (Table 1). Beyond numerical gains, the adaptive spline provides greater interpretability through progression curves that align with clinical patterns reported in Alzheimer's research. This underscores adaptive knot selection as an effective strategy for improving accuracy while yielding more realistic representations of disease trajectories. The work extends prior methodological findings on the importance of knot selection by demonstrating the effectiveness of an adaptive spline framework in the longitudinal neurodegenerative setting, with both numerical and visual validation. A principal limitation of this study is the use of simulated data, which, although biologically informed, cannot fully capture the complexity of real-world Alzheimer's progression. Future research will apply the proposed framework to real longitudinal datasets such as OASIS and ADNI, with a particular focus on mixed-effects GAM formulations to account for subject-specific variability. Further extensions will explore the integration of multimodal biomarkers, including MRI-derived measures and clinical covariates, to enhance both predictive performance and clinical relevance.

REFERENCES

[1] Z. Xiaopeng *et al.*, "Global Burden of Alzheimer's disease and other dementias in adults aged 65 years and older, 1991–2021: population-based study," *Front. Public Heal.*, vol. 13, 2025, doi: 10.3389/fpubh.2025.1585711.

[2] A. P. Porsteinsson, R. S. Isaacson, S. Knox, M. N. Sabbagh, and I. Rubino, "Diagnosis of Early Alzheimer's Disease: Clinical Practice in 2021," *J. Prev. Alzheimer's Dis.*, vol. 8, no. 3, pp. 371–386, 2021, doi: 10.14283/jpad.2021.23.

[3] I. Nagarajan and G. G. Lakshmi Priya, "A comprehensive review on early detection of Alzheimer's disease using various deep learning techniques," *Front. Comput. Sci.*, vol. 6, 2024, doi: 10.3389/fcomp.2024.1404494.

[4] S. Golriz Khatami, C. Robinson, C. Birkenbihl, D. Domingo-Fernández, C. T. Hoyt, and M. Hofmann-Apitius, "Challenges of Integrative Disease Modeling in Alzheimer's Disease," *Front. Mol. Biosci.*, vol. 6, no. January, pp. 1–13, 2020, doi: 10.3389/fmolb.2019.00158.

[5] D. Chen, M. Zhang, H. Han, Y. Wen, and H. Yu, "Reflections on dynamic prediction of Alzheimer's disease: advancements in modeling longitudinal outcomes and time-to-event data," *BMC Med. Res. Methodol.*, vol. 25, no. 1, 2025, doi: 10.1186/s12874-025-02618-x.

[6] A. Perperoglou, W. Sauerbrei, M. Abrahamowicz, and M. Schmid, "A review of spline function procedures in R," *BMC Med. Res. Methodol.*, vol. 19, no. 1, pp. 1–16, 2019, doi: 10.1186/s12874-019-0666-3.

[7] A. Araveeporn, "The Estimating Parameter and Number of Knots for Nonparametric Regression Methods in Modelling Time Series Data," *Modelling*, vol. 5, no. 4, pp. 1413–1434, 2024, doi: 10.3390/modelling5040073.

[8] V. Goepp, O. Bouaziz, and G. Nuel, "Spline regression with automatic knot selection," *Comput. Stat. Data Anal.*, vol. 202, no. August 2024, p. 108043, 2025, doi: 10.1016/j.csda.2024.108043.

[9] A. Fajry, A. Barra, D. Retno, and S. Saputro, "Knot Optimization For Bi-Response Spline Nonparametric Regression With Generalized Cross-Validation (GCV)," *BAREKENG J. Math. Its Appl.*, vol. 19, no. 1, pp. 271–280, 2025.

[10] T. Handayani, S. Sifriyani, and A. T. R. Dani, "Nonparametric Spline Truncated Regression with Knot Point Selection Method Generalized Cross Validation and Unbiased Risk," *JTAM (Jurnal Teor. dan Apl. Mat.*, vol. 7, no. 3, p. 848, 2023, doi: 10.31764/jtam.v7i3.14034.

[11] S. Yu, G. Wang, L. Wang, and L. Yang, "Multivariate spline estimation and inference for image-on-scalar regression," *Stat. Sin.*, vol. 31, no. 3, pp. 1463–1487, 2021, doi: 10.5705/ss.202019.0188.

[12] H. Zhao, Y. Zhang, X. Zhao, and Z. Yu, "A nonparametric regression model for panel count data analysis," *Stat. Sin.*, vol. 29, no. 2, pp. 809–826, 2019, doi: 10.5705/ss.202016.0534.

[13] M. C. Donohue *et al.*, "Natural cubic splines for the analysis of Alzheimer's clinical trials," *Pharm. Stat.*, vol. 22, no. 3, pp. 508–519, 2023, doi: 10.1002/pst.2285.

[14] D. A. Jenkins, M. Sperrin, G. P. Martin, and N. Peek, "Dynamic models to predict health outcomes: current status and methodological challenges," *Diagnostic Progn. Res.*, vol. 2, no. 1, p. 23, 2018, doi: https://doi.org/10.1186/s41512-018-0045-2.

[15] A. Cascarano *et al.*, "Machine and deep learning for longitudinal biomedical data: a review of methods and applications," *Artif. Intell. Rev.*, vol. 56, no. Suppl 2, pp. 1711–1771, 2023, doi: https://doi.org/10.1007/s10462-023-10561-w.

[16] J. G. Ibrahim and G. Molenberghs, "Missing data methods in longitudinal studies: a review," *Test*, vol. 18, no. 1, pp. 1–43, 2009, doi: https://doi.org/10.1007/s11749-009-0138-x.

[17] M. E. Griswold *et al.*, “Reflection on modern methods: shared-parameter models for longitudinal studies with missing data,” *Int. J. Epidemiol.*, vol. 50, no. 4, pp. 1384–1393, 2021, doi: https://doi.org/10.1093/ije/dyab086.
[18] G. Chen *et al.*, “Beyond linearity in neuroimaging: Capturing nonlinear relationships with application to longitudinal studies,” *Neuroimage*, vol. 176, no. 3, pp. 139–148, 2019, doi: 10.1016/j.neuroimage.2021.117891.Beyond.
[19] G. Chen *et al.*, “Beyond linearity in neuroimaging: Capturing nonlinear relationships with application to longitudinal studies,” *Neuroimage*, vol. 233, no. 2, p. 117891, 2021, doi: https://doi.org/10.1016/j.neuroimage.2021.117891.
[20] S. C. Villeneuve *et al.*, “Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer’s disease,” *Alzheimer’s Dement. Transl. Res. Clin. Interv.*, vol. 5, pp. 553–562, 2019, doi: 10.1016/j.trci.2019.08.009.
[21] S.-C. Villeneuve *et al.*, “Latent class analysis identifies functional decline with Amsterdam IADL in preclinical Alzheimer’s disease,” *Alzheimer’s Dement. Transl. Res. Clin. Interv.*, vol. 5, pp. 553–562, 2019, doi: https://doi.org/10.1016/j.trci.2019.08.009.
[22] A. Wearn, L. L. Raket, D. L. Collins, and R. N. Spreng, “Longitudinal changes in hippocampal texture from healthy aging to Alzheimer’s disease,” *Brain Commun.*, vol. 5, no. 4, pp. 1055–1064, 2023, doi: 10.1093/braincomms/fcad195.
[23] D. S. Marcus *et al.*, “Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI Data in Young , Middle Aged , Nondemented , and Demented Older Adults Citation Nondemented , and Demented Older Adults,” *Digit. Access to Scholarsh. Harvard*, 2007, doi: 10.1162/jocn.2007.19.9.1498.Published.
[24] D. S. Marcus, A. F. Fotenos, J. G. Csernansky, J. C. Morris, and R. L. Buckner, “Open Access Series of Imaging Studies (OASIS): Longitudinal MRI Data in Nondemented and Demented Older Adults,” *J. Cogn. Neuoscience*, vol. 22, no. 12, 2010, doi: 10.1162/jocn.2009.21407.
[25] C. R. J. Jr *et al.*, “NIA-AA Research Framework Toward a biological definition of Alzheimer s disease.pdf,” 2018.
[26] F. Pedregosa, R. Weiss, and M. Brucher, “Scikit-learn : Machine Learning in Python,” *J. ofMachine Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
[27] M. J. Azur, E. A. Stuart, C. Frangakis, P. J. Leaf, and D. C. Washington, “Multiple imputation by chained equations : what is it and how does it work ?,” *Int. J. Methods Psychiatr. Res.*, vol. 20, no. 1, pp. 40–49, 2011, doi: 10.1002/mpr.