




Cite this: *Soft Matter*, 2025,
21, 729

Role of data-driven regional growth model in shaping brain folding patterns†

Jixin Hou,^a Zhengwang Wu,^b Xianyan Chen,^c Li Wang,^b Dajiang Zhu,^d
Tianming Liu,^e Gang Li^{*b} and Xianqiao Wang  ^{*a}

The surface morphology of the developing mammalian brain is crucial for understanding brain function and dysfunction. Computational modeling offers valuable insights into the underlying mechanisms for early brain folding. Recent findings indicate significant regional variations in brain tissue growth, while the role of these variations in cortical development remains unclear. In this study, we explored how regional cortical growth affects brain folding patterns using computational simulation. We first developed growth models for typical cortical regions using machine learning (ML)-assisted symbolic regression, based on longitudinal real surface expansion and cortical thickness data from prenatal and infant brains derived from over 1000 MRI scans of 735 pediatric subjects with ages ranging from 29 postmenstrual weeks to 2 years of age. These models were subsequently integrated into computational software to simulate cortical development with anatomically realistic geometric models. We comprehensively quantified the resulting folding patterns using multiple metrics such as mean curvature, sulcal depth, and gyrification index. Our results demonstrate that regional growth models generate complex brain folding patterns that more closely match actual brains structures, both quantitatively and qualitatively, compared to conventional uniform growth models. Growth magnitude plays a dominant role in shaping folding patterns, while growth trajectory has a minor influence. Moreover, multi-region models better capture the intricacies of brain folding than single-region models. Our results underscore the necessity and importance of incorporating regional growth heterogeneity into brain folding simulations, which could enhance early diagnosis and treatment of cortical malformations and neurodevelopmental disorders such as cerebral palsy and autism.

Received 10th October 2024,
Accepted 29th December 2024

DOI: 10.1039/d4sm01194e

rsc.li/soft-matter-journal

1. Introduction

As the central regulator of physiological activities, the human brain exhibits an intricate tissue structure marked by pronounced heterogeneity. During early development (16–40 postmenstrual weeks), the brain undergoes considerable expansion in its volume and cortical surface area, accompanied by the noticeable observation of cortical folds, with the convex peak known as gyri and the concave valley as sulci. Typically, cortical

folding begins with primary folds, which are highly conserved across individuals,¹ and gradually progresses to more complex secondary and tertiary folds that exhibit significant individual variability.² Emerging evidence links abnormal cortical folding patterns to various neurodevelopmental disorders such as autism,³ epilepsy,⁴ and schizophrenia,⁵ as well as cortical malformations like lissencephaly, pachygyria, and polymicrogyria.⁶ Consequently, a comprehensive understanding of the development of cortical folding is crucial for early detection and treatment of cognitive impairments and neurodevelopmental disorders.

The development of cortical folding involves highly complex spatiotemporal patterns influenced by multiple physiological factors,⁷ such as cranial constraints,⁸ axon maturation,⁹ cerebrospinal fluid,¹⁰ genetic expression,¹¹ etc. Though the precise mechanisms underlying cortical fold formation remains elusive, increasing studies suggest that mechanics play a crucial role in modulating this process.¹² Brain tissue development occurs within a complex physical environment comprising the skull, meninges, and cerebrospinal fluid. While external forces such as constraints from the skull and cerebrospinal fluid

^a School of Environmental, Civil, Agricultural and Mechanical Engineering, College of Engineering, University of Georgia, Athens, GA 30602, USA. E-mail: xqwang@uga.edu

^b Department of Radiology and Biomedical Research Imaging Center, The University of North Carolina at Chapel Hill, NC 27599, USA. E-mail: gang_li@med.unc.edu

^c Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, GA 30602, USA

^d Department of Computer Science and Engineering, The University of Texas at Arlington, Arlington, TX 76019, USA

^e School of Computing, The University of Georgia, Athens, GA 30602, USA

† Electronic supplementary information (ESI) available. See DOI: <https://doi.org/10.1039/d4sm01194e>



pressure are expected to affect cortical folding patterns, experimental observations indicate that the internal forces exerted within brain tissue are more dominant at the onset of folding.¹³ In line with these findings, Essen¹⁴ firstly proposed the “axon tension” theory, which suggests that axonal tension pulls gyral walls together, leading to formation of cortical folds. In contrast, Nie *et al.*¹⁵ proposed that pushing force from axonal growth could generate folding shapes similar to those observed in real cortical surface. However, despite experimental evidence of forces acting on axons, their magnitude and direction appear insufficient to pull or push brain tissues into folds.¹⁶ Conversely, the differential growth theory, which has garnered more experimental support,^{17,18} posits that the mismatch in growth ratios between two-layered brain structure triggers the bulking of cortical surface.¹⁹ Specifically, the fast-growing outer layer (gray matter), under connectivity constraint of slow-growing inner layer (white matter), generates excessive compressive forces that initiate cortical folding. Therefore, it is essential to investigate how varying tissue growth influences the initiation and regulation of cortical folding patterns.

Computational simulation, such as finite element method (FEM), have emerged as a powerful tool for modeling the spatiotemporally complex development of cortical folding patterns, due to its superiority in reproducibility and cost-effectiveness compared to experimental and longitudinal imaging methods.²⁰ To replicate folding patterns, most studies have assumed uniform growth patterns such as isotropic growth and purely tangential growth across growing tissues.^{21–25} While these approaches have greatly advanced our understanding of cortical folding mechanisms, their effectiveness diminishes when comparing the simulated patterns to actual brain structures, particularly in brain-wide simulations.^{26,27} This discrepancy arises because uniform growth assumptions conflict with the heterogeneous nature of growth as observed in experimental and imaging investigations.^{28,29} Recent studies, including our own study,³⁰ have shown that the cortical surface can be parcellated into 18 distinct regions, each exhibiting significantly different area expansion ratios from the third trimester of pregnancy to

the first two years post-birth, as illustrated in Fig. 1. Similar parcellation based on heterogeneous development of cortical thickness have also been reported.³¹ These observations underscore the pronounced spatiotemporal heterogeneity in cortical tissue growth.

Despite its significance, the effect of growth heterogeneity on cortical folding patterns has been largely overlooked in computational studies. Budday and Steinmann³² were among the first to investigate this effect by incorporating manually designed sinusoidal variations in growth distributions during folding simulations based on idealized models. Their findings revealed that regional variations in cortical growth primarily influence secondary instabilities, contributing to the irregularity of folding patterns. Wang *et al.*³³ further explored this by simulating cortical folding evolution on theoretical brain models with regionally heterogeneous growth, linking the growth ratio to local curvature. Their results demonstrate the role of regional growth in shaping complex cortical folding patterns, characterized by fast-growing gyri and slow-growing sulci. Similar conclusions were drawn by Zhang *et al.*,³⁴ who reported that gyral convex patterns consistently emerge in cortical regions with faster growth speed. While these studies have enhanced our understanding of the critical role of heterogeneous growth in cortical folding development, they have primarily focused on spatial heterogeneity and relied on idealized geometric models with manually defined growth patterns. These approaches deviate significantly from real brain scenarios. A comprehensive study examining the effects of realistic heterogeneous growth on cortical folding patterns in models that closely resemble real brains is still lacking.

In this study, we aim to explore how anatomically realistic regional growth affects the simulated cortical folding patterns on geometric models derived from real brains. We first employed ML-assisted symbolic regression to generate regional growth models using longitudinal surface area and cortical thickness data, including over 1000 infant magnetic resonance imaging (MRI) scans from developing prenatal and neonatal human brains. These predicted growth models were then

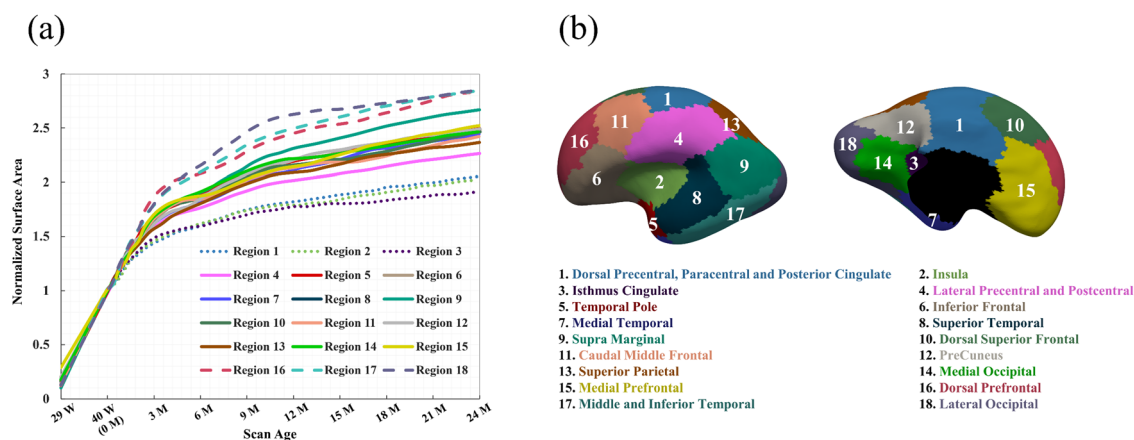


Fig. 1 Developmental trajectory of surface area in each region. (a) Surface area is normalized based on the 40 postmenstrual weeks. W and M represent weeks and months, respectively; (b) developmental regionalization map of 18 regions.



integrated into ABAQUS to simulate mechanical folding development using tangential differential growth theory, applied to regional geometric models constructed from initial brain surfaces. The spatiotemporal folding patterns were analyzed and compared with those observed in real brains, both qualitatively and quantitatively. Additionally, we conducted an in-depth analysis by decomposing the growth's value and trajectory to assess their individual impacts on cortical folding patterns. Our results highlight the importance of considering growth heterogeneity in simulations of brain folding development. This approach could potentially contribute to early diagnosis of cortical malformations such as lissencephaly and polymicrogyria and improve treatments for neurodevelopmental disorders like epilepsy and autism.

2. Methods

2.1. Biomechanics in modeling brain folding

Differential growth theory indicates that cortical folding during brain development arises due to disparities in growth ratio between the faster-growing cortical layer and the slower-growing white matter layer. To investigate the effect of regional growth on folding evolution in the developing brain, we conducted nonlinear finite element simulations grounded in continuum mechanics theory and the finite growth assumption. In describing deformation kinematics, we introduce a one-to-one mapping denoted as $\mathbf{x} = \boldsymbol{\varphi}(\mathbf{X})$, which carries a material particle initially positioned at \mathbf{X} in the reference configuration \mathcal{B}_0 to its new position \mathbf{x} in the current configuration \mathcal{B}_t . Leveraging the theory of multiplicative decomposition proposed by Rodriguez *et al.*,³⁵ we decompose the deformation gradient \mathbf{F} and the Jacobian J into elastic and growth components:

$$\mathbf{F} = \nabla_{\mathbf{x}}\boldsymbol{\varphi} = \mathbf{F}^e \cdot \mathbf{F}^g \quad \text{and} \quad J = \det \mathbf{F} = J^e \cdot J^g. \quad (1)$$

We assume that brain tissue growth follows an orthotropic manner,

$$\mathbf{F}^g = g_t \mathbf{I} + (g_r - g_t) \hat{\mathbf{n}} \otimes \hat{\mathbf{n}} \quad \text{and} \quad J^g = \det \mathbf{F}^g, \quad (2)$$

where g_t and g_r are the relative growth ratio defined in the tangential (in-plane) and radial (out-of-plane) directions, respectively, of which the expressions ($g_t = g_t(t)$, $g_r = g_r(t)$) are derived from the symbolic regression predictions, as discussed in the following section. g_t equating to g_r indicates an isotropic growth scenario, while $g_r = 1$ denotes tangential growth. The vector $\hat{\mathbf{n}}$ is the local surface normal, oriented positively in the out-of-plane direction. Growth kinematics alone generate a stress-free state, which must be balanced by the elastic deformation gradient introduced by material confinement to avoid incompatible deformation.³⁶

We characterized the constitutive behavior of the brain tissue through the following neo-Hookean strain energy function, parameterized exclusively in terms of the elastic components of deformation gradient \mathbf{F}^e and its Jacobian J^e ,

$$\Psi(\mathbf{F}^e) = \frac{1}{2} \lambda \ln^2(J^e) + \frac{1}{2} \mu (\mathbf{F}^e : \mathbf{F}^e - 2 \ln J^e - 3), \quad (3)$$

where λ and μ are the Lamé constants. In accordance with thermodynamic principles, the first Piola–Kirchhoff stress \mathbf{P} works conjugate with the deformation gradient \mathbf{F} , particularly \mathbf{F}^e , since only the elastic part of deformation produces stress,

$$\mathbf{P} = \frac{\partial \Psi}{\partial \mathbf{F}} = \frac{\partial \Psi(\mathbf{F}^e)}{\partial \mathbf{F}^e} : \frac{\partial \mathbf{F}^e}{\partial \mathbf{F}} = \mathbf{P}^e \cdot \mathbf{F}^{e-T}. \quad (4)$$

In the absence of body force, the balance of linear momentum reduces to the vanishing divergence of the first Piola–Kirchhoff stress \mathbf{P} :

$$\text{Div } \mathbf{P} = 0 \quad (5)$$

2.2. Symbolic regression for discovering growth models

Inspired by Darwinian principles of natural selection, symbolic regression autonomously uncovers mathematical relationships exclusively from provided data without requiring prior knowledge, thereby significantly enhancing the interpretability and flexibility of the model discovery process.³⁷ It has demonstrated promising applications in model characterization^{38–40} and parameter calibration.⁴¹ Symbolic regression operates through a process known as genetic programming (GP). During GP execution, functional expressions are efficiently formatted using a binary-tree structure, which consists of nodes and branches.

A complete tree structure, as illustrated in Fig. 2a, involves variables, mathematical operators (either unary or binary), and constants. The evolution process experiences the genetic operations of evaluation, selection, mutation, and crossover, while the latter two are essential for updating the tree structure. Specifically, the mutation operation amplifies population's genetic diversity by randomly altering some nodes in an expression tree, as exemplified in Fig. 2b, where a new offspring is generated by substituting the exponential operator (exp) with the hyperbolic tangent (tanh). Conversely, the crossover operation allows the algorithm to create new offspring by combining building blocks from different parent candidates, as demonstrated in Fig. 2c. This iterative process of evaluation, selection, mutation, and crossover continues until the optimal expression is obtained or the maximum number of generations is reached, whichever is reached first.³⁷

In this study, we employed symbolic regression to discover appropriate growth models for the human brain cortex. The raw data includes the measured surface area of each parcellated region in the developing brain cortex along with the corresponding gestational ages, as shown in Fig. 3a.

To facilitate computational implementation, we first converted the data into unitless growth ratio $g(t)$ and virtual time t using the following formulas:

$$g_t(t) = \sqrt{\frac{S(t)}{S_0}}, \quad g_r(t) = \sqrt{\frac{T(t)}{T_0}}, \quad \text{and} \quad t = \frac{\text{GA}(t) - \text{GA}_0}{\text{GA}_{\max} - \text{GA}_{\min}}, \quad (6)$$

where $g_t(t)$ and $g_r(t)$ are tangential and radial growth ratio, respectively. S_0 and T_0 denotes the surface area and cortical



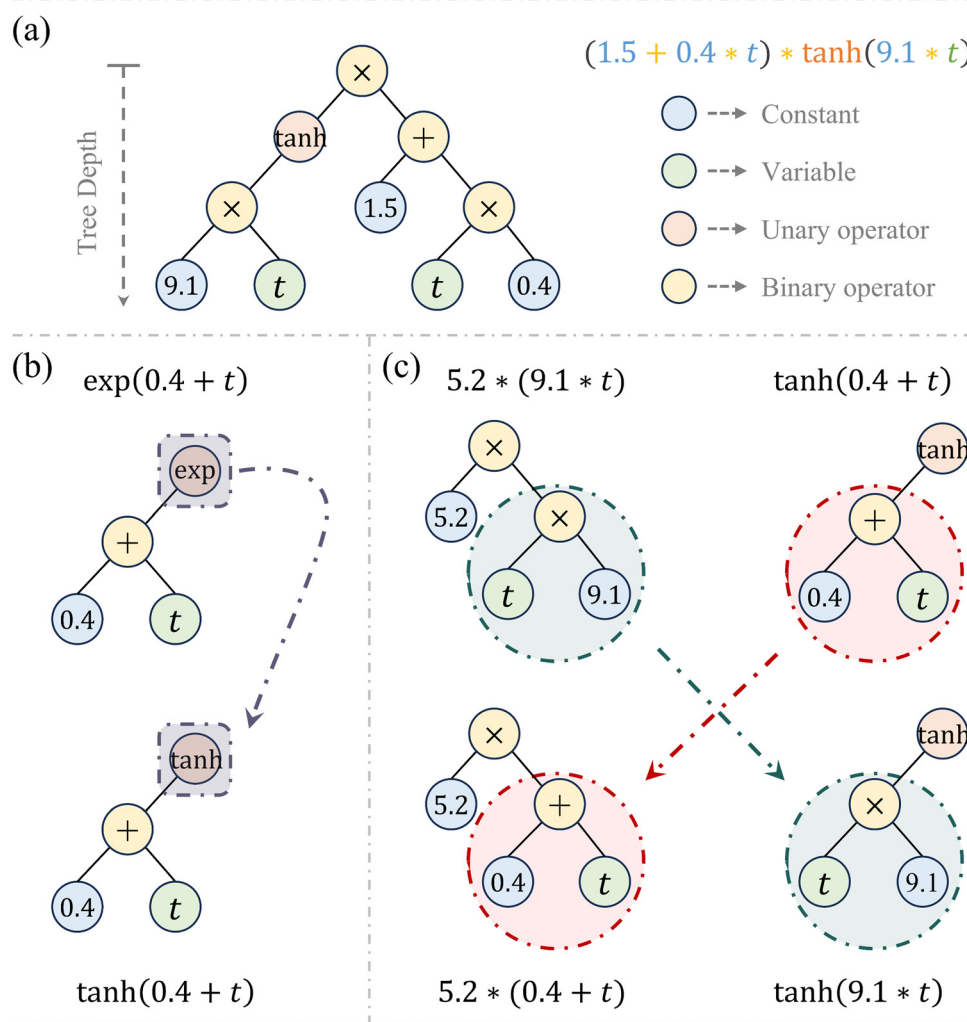


Fig. 2 Structure and operations of expression tree. (a) Representation of tree structure for an algebraic expression; (b) an example of the mutation operation; (c) an example of the crossover operation between two parent trees.

thickness measured at the initial gestational age, GA_0 (29 postmenstrual weeks), respectively. GA_{\max} and GA_{\min} corresponds to the maximum and minimal gestational age within the measured data range, herein, their values are 29 postmenstrual weeks and 24 postnatal months of age, respectively. Through the above conversion, the value of t ranges from 0 to 1, which serves as the input for the symbolic regression algorithm to find the optimal growth model $g(t)$ for each region, as illustrated in Fig. 3a.

All symbolic regression analyses were performed using PYSR,⁴² a powerful open-source package developed alongside the Julia library SymbolicRegression.jl. During the training of symbolic regression, the binary operators are restricted to addition (+), multiplication (*), and polynomial (t^n), while the unary operators are limited to the exponential (exp), hyperbolic tangent (tanh), square (t^2), cube (t^3), and natural logarithmic (ln) functions. Given the balance between computational cost and model performance, the training time is limited to 30 minutes or a maximum of 1000 iterations, whichever is reached first, as our observations indicate that most qualified

models could be identified within this timeframe. The maximum depth of the expression tree is set to 10, and the maximum complexity is constrained to 100. Notably, the nested behavior of the exponential, hyperbolic tangent, and logarithmic functions is forbidden, while the square and cube operators are restricted to occur, if desired, only once inside the exponential, hyperbolic tangent, and logarithmic functions. We adopt the default criterion ("best") to guide the model selection process. For each algorithm, the training is repeated at least three times, and favorable candidates are selected as the target models. All training was performed on a Legion PC equipped with a six-core Intel Core i7-8750H 2.2 GHz CPU, 4 GB NVIDIA GTX 1050Ti GPU, and 24 GB of memory.

2.3. Computational modeling of a developing brain

The identified regional growth models (both tangential and radial) were used to construct the growth tensor F^g , which can be applied to simulate folding evolution using the FEM, as shown in Fig. 3b. The simulation results were then compared



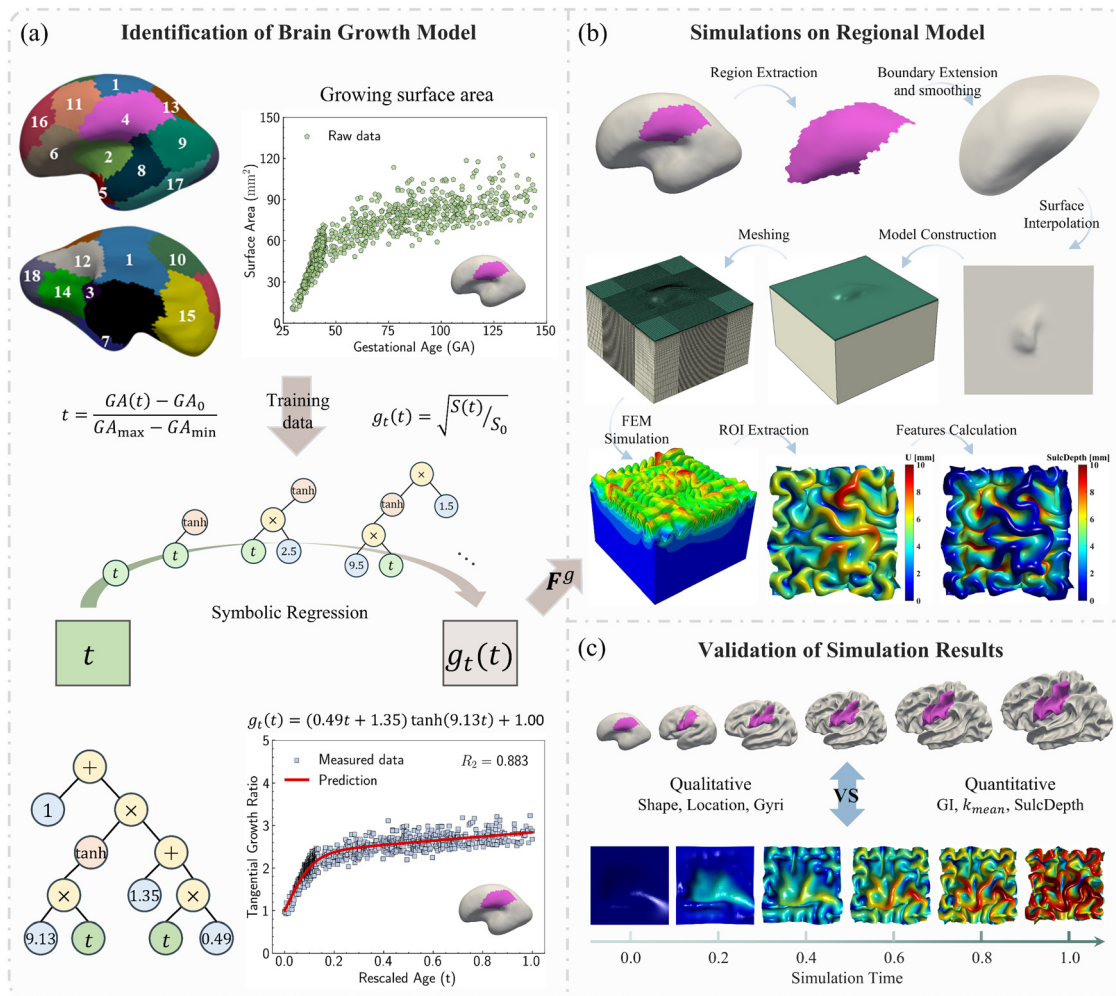


Fig. 3 Schematic diagram of this research. (a) Symbolic regression algorithms used to identify the mathematical growth models for each region based on regional developmental surface area data from 29 postmenstrual weeks to 24 postnatal months of age, exemplified with region 4; (b) construction process of finite element model for each region, ROI means region of interest. (c) Qualitative and quantitative comparison between the simulation results and real human brains measures along the developmental timeline.

with the brain imaging data for model validation, as illustrated in Fig. 3c. To simulate the folding evolution of the developing brain, we constructed a three-dimensional double-layer patch model based on the geometries of a human brain at 29 postmenstrual weeks, as depicted in Fig. 3b. Initially, the regional brain inner surface (the interface between the gray matter and white matter) was extracted using the parcellation map provided by Huang *et al.*³⁰ The extracted surface was first extended by 2–5 mm along the boundary's local curvature. Laplacian smoothing was then applied to the boundary area using a smoothing parameter of 0.2, with the number of iterations ranging from 15 to 22 across all extracted surfaces. During smoothing, the intermediate area remains fixed to preserve the integrity of the initial geometry (see Fig. S1 in the ESI†). Subsequently, we interpolated the extended surface with a flat plane of 80 mm × 80 mm and merged these two surfaces using Boolean operations. The connecting areas were further smoothed to ensure a natural transition of the curvature. This extension and interpolation ensured that the model's

dimensions were large enough compared to the wavelength of folded patterns observed in experiments, thereby preventing boundary effects.²⁵ Additionally, the squared boundaries significantly simplified the prescription of boundary conditions during modeling. We then shifted the interpolated surface upwards by 2 mm to form the initial cortical layer and extended the squared boundary downwards by 50 mm to generate the initial white matter layer. This design was based on experimental observations in neonatal human brains, which indicate that the cerebral cortex is a thin layer with a thickness of 2–3.5 mm, while the core has a much greater thickness of around 50 mm.⁴³ Consequently, the base model's dimensions were approximately 80 mm × 80 mm × 50 mm (excluding cortical thickness), as illustrated in Fig. 4a, where $l = 80$ mm, $h_s = 50$ mm, and $h_c = 2$ mm.

All simulations were performed using the commercial software ABAQUS (Dassault Systems, Paris, France).⁴⁴ Dynamic-explicit solver was employed due to its superior performance in solving nonlinear, dynamic, and larger deformation



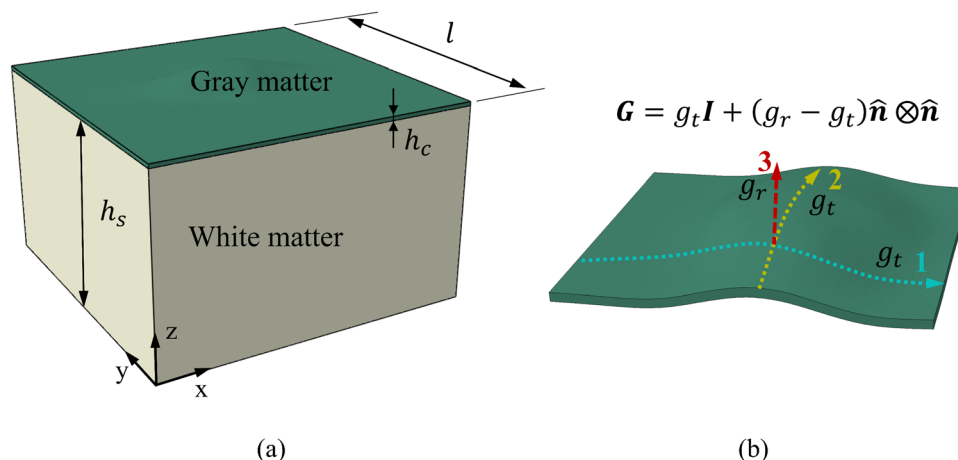


Fig. 4 Geometric model for modeling regional brain growth. (a) Initial status of a FEM model, h_c and h_s represent the layer thickness of gray matter and white matter, respectively; (b) orthogonal growth patterns, g_t and g_r indicate the tangential growth ratio within the plane and radial growth ratio in the thickness direction, respectively.

problems.^{45,46} Both the gray and white matters were modeled as incompressible neo-Hookean materials, with elastic stiffness values of 0.31 kPa for the cortical and 0.45 kPa for the white matter layer.⁴⁷ For detailed material parameter settings, please refer to the Table S1 in the ESI.[†] Orthotropic growth was defined for the cortical layer, while isotropic growth was applied to the white matter layer. In our modeling approach, growth was simulated using thermal expansion, considering the analogy between the volumetric growth and the thermal expansion.⁴⁸ The expansion ratio $\alpha(t)$ correlates to the growth ratio as $\alpha(t) = g(t) - 1$. Specifically, for the white matter layer, the expansion ratio was defined as $\alpha(t) = 0$, while for the cortical layer, $\alpha(t) = g_r(t) - 1$ was applied to out-of-plane growth and $\alpha(t) = g_t(t) - 1$ to in-plane growth, as shown in Fig. 4b. The customized growth models ($g_r(t)$ and $g_t(t)$), derived from symbolic regression, were implemented into the finite element algorithm through a user-defined subroutine VUEXPAN. Symmetric boundary conditions were prescribed on the four sides of the model and the bottom surface of the white matter layer was fixed. Free boundary conditions were applied to the top surface of the cortical layer, accompanied by a self-contact constraint to prevent self-penetration. The total simulation time was set to 1 s. To avoid computational instabilities, the maximum time step was determined as $dt = 0.05a\sqrt{\rho/K}$, where a is the average mesh size, ρ is the mass density, K is bulk modulus.⁴⁹ Temperature variation was applied using a sigmoidal smooth step function.

Structural meshing with the element type C3D8R was conducted for both the cortical and white matter layer. To determine the appropriate mesh size, we conducted a mesh sensitivity analysis with mesh size ranging from 0.3 mm to 0.8 mm (ESI,[†] Fig. S2 and S3). Based on the mesh convergence analysis—where simulation results with the coarsest mesh closely matched those of the finest mesh—we selected a mesh size of 0.5 mm for all models. This results in 84 700 elements for the cortical layer and 278 300 elements for the white matter layer. During the simulations, we recorded the coordinates and

displacements of the region of interest (ROI) for each frame. The ROI was defined as the smallest square area encompassing the extracted brain surface region. Although this definition introduces some redundant areas, which potentially biases the quantitative measurements, it serves our primary goal: to compare the effectiveness of the regional growth model with the widely used isotropic growth theory. The inclusion of these redundant areas does not significantly impact this comparative analysis. Additionally, defining the ROI in this manner simplifies the partitioning process in ABAQUS and facilitates area reconstructions during postprocessing. All simulations were performed on a Dell workstation equipped with a 16-core Intel(R) Xeon(R) CPU E5-2687 W@3.1 GHz, and 64 GB of memory.

2.4. Postprocessing and quantitative metrics

After the simulations, the recorded coordinates and displacements were first extracted from the result file using Python and subsequently imported into MATLAB to reconstruct the deformed surface. During reconstruction, the surface was interpolated five times to generate a sufficiently smooth surface, and the original quadrilateral surface mesh was transformed into a triangular mesh, facilitating the calculation of quantitative features such as curvatures, gyrification index, and sulcal depth in MATLAB.

2.4.1. Curvatures. The curvature of a surface describes the degree to which it deviates from being flat at a given point. Normal curvature is defined as the inverse of the radius of the best-approximated curve from a surface normal slice in a given direction. Considering all directions, we obtain the curvature matrix, typically represented by the Weingarten matrix. Its principal decomposition gives the principal curvatures, which correspond to the maximum and minimum values of the surface's normal curvature in different directions (k_1 and k_2). The average of two principal curvatures denotes the mean curvature ($k_H = (k_1 + k_2)/2$), while the product of the principal curvatures yields the Gaussian curvature ($k_G = k_1 \cdot k_2$). In this study, we focused on the mean curvature due to its extensive application in brain cortical folding analysis.^{49–51}



The curvatures of the triangular mesh were calculated using the method introduced by Meyer *et al.*,⁵² where finite volume discretization was employed to estimate the local integral of mean curvature over the normal areas of triangular faces associated with each point. If the surrounding faces are obtuse triangles, the barycentric area was calculated; otherwise, the Voronoi area was used. However, the calculated mean curvature is dependent on the geometry's shape or size, meaning its magnitude varies with brain scales. To address this, we further introduced a non-dimensional measure of mean curvature using the method provided by Balouchzadeh *et al.*,⁵³ where the mean curvature is multiplied by a characteristic length l_c ,

$$l_c = \sqrt{A_t/4\pi} \Rightarrow k_H^* = k_H \cdot l_c, \quad (7)$$

where A_t is the surface area. The dimensionless mean curvature was calculated for each point to provide a qualitative representation, while the absolute value of the dimensionless mean curvature was averaged across all model points for quantitative comparison. In the remainder of the manuscript, we use the term "curvature" to refer to "dimensionless mean curvature" for clarity.

2.4.2. Gyrfication index. To quantitatively describe the folding complexity of the deformed brain surface, we introduced a global folding metric: the three-dimensional gyrfication index (GI). The GI is defined as the ratio of the total cortical surface area to the area of convex hull that completely encloses the convoluted surface,⁸

$$GI = \frac{\text{area of cortical surface}}{\text{area of convex hull}}. \quad (8)$$

To calculate the GI, we first defined a fully enclosed convex hull comprising all points of the deformed cortical surface, then we discretized and filtered this surface to ensure it completely encloses the deformed surface with minimum surface area. Finally, we measured the area of discrete convex hull, which serves as the denominator in the GI calculation.

2.4.3. Sulcal depth. Sulcal depth (SulcDepth) is another quantitative measure capable of reflecting the extent of the folding in brain regions. Although Numerous methods have been suggested for computing sulcal depth,^{51,54} a well-defined computation remains elusive. In this study, we adopted the approach introduced by Wang *et al.*,⁴⁹ calculating SulcDepth as the distance between the deformed mesh surface and its convex hull, which was previously defined in calculating the GI. Specifically, for each vertex on the deformed surface, we first determined its projection point on each discrete triangular surface of the convex hull. Subsequently, we computed the distance between the vertex and its corresponding projection point, with the shortest distance serving as the sulcal depth for that vertex. Given the convex nature of the enclosed hull, the shortest distance always exists between the vertex and a consistent piece of the convex hull, as demonstrated in the Result section. SulcDepth was calculated for each vertex for qualitative representation, and the values were averaged across all model points to ensure quantitative comparisons.

2.4.4. Evaluation metrics. To evaluate the accuracy of simulation results for the aforementioned quantitative features, including curvatures, gyrfication index, and sulcal depth, we introduced the following statistical metrics: R^2 , mean absolute percentage error (MAPE), and Pearson correlation coefficient (r). Here, R^2 measures the goodness-of-fit of the simulation results, formulated as $R^2 = 1 - \sum_{i=1}^N (y_i - \hat{y}_i)^2 / (y_i - \bar{y})^2$, where N denotes the number of data points, y_i represents the actual feature values measured from real brain data, \bar{y} is the mean of actual values, and \hat{y}_i is the predicted value from the simulation. An R^2 value close to 1 indicates a strong agreement between the simulation results and real brain imaging data. MAPE, expressed as a percentage, quantifies the average absolute error between predicted (\hat{y}_i) and actual values (y_i). It is calculated using the formula:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (9)$$

Moreover, we introduced the Pearson correlation coefficient (r) to assess the strength and direction of the linear relationship between the simulation results and real brain measures, with its formula as:

$$r = \frac{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (\hat{y}_i - \bar{\hat{y}})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (10)$$

where $\bar{\hat{y}}$ is the mean of predicted values. All these evaluation metrics were calculated in Python using the sklearn and scipy libraries. Noted, due to the discrepancy in the number of data points between the simulation results and the actual data, we applied an interpolation method to resample the extracted simulation data, ensuring a consistent data size before computing the evaluation metrics.

2.5. Brain imaging data

In this study, we aimed to conduct symbolic regression to identify tangential growth models by measuring brain surface area data from two high-quality, publicly available imaging datasets focused on early brain development: the developing Human Connectome Project (dHCP) and the Baby Connectome Project (BCP). The dHCP dataset includes 549 MRI scans from 500 healthy neonates (279 males, 221 females) aged between 29 and 45 postmenstrual weeks.⁵⁵ These MR images were obtained using a Philips 3T scanner with a 32-channel neonate-dedicated head coil at St. Thomas Hospital, London, UK.⁵⁶ The BCP dataset comprises 488 longitudinal MRI scans from 235 healthy, term-born infants (108 males, 127 females) scanned between 41 and 144 postmenstrual weeks. These images were collected using 3T Siemens Prisma MRI scanners equipped with Siemens 32-channel head coils at The University of North Carolina at Chapel Hill and University of Minnesota. To investigate the developmental regionalization of cortical surface area expansion during pregnancy and infancy, a data-driven method known as non-negative matrix factorization (NMF) was utilized to partition the



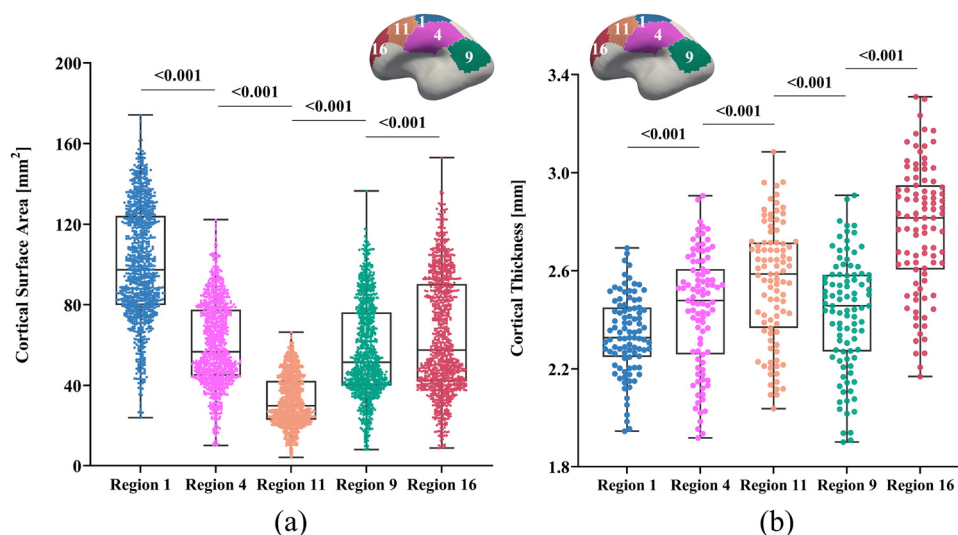


Fig. 5 Raw data of cortical surface area and thickness for five selected regions. Analysis reveals significant differences in the development of surface area (a) and thickness (b) among these regions, with all p -values less than 0.001. Insets show spatial locations of regions on the cortical surface. Dots denote MRI scans collected from 29 postmenstrual weeks to 2 years of age.

cortical surface into distinct regions by clustering cortical vertices with similar developmental patterns. For more details, please refer to our recent work.³⁰

The radial growth model was characterized using thickness data measured in the developing neonatal brain.³¹ It is important to note that the radial growth dataset consisted exclusively of postnatal measurement. Consequently, we could not perform symbolic regression on a continuous dataset spanning from 29 to 136 postmenstrual weeks, as we did for the tangential growth model. To address this limitation, we assumed a linear increase in cortical thickness during the perinatal period (29–40 postmenstrual weeks) based on brain imaging analyses^{57,58} and applied linear extrapolation to ensure a smooth transition. However, it should be mentioned that there is no clear consensus on the exact developmental trajectory of cortical thickness during the perinatal stage. For example, Demirci and Holland⁵⁹ recently reported a weak linear correlation between cortical thickness and scan ages during 29–43 postmenstrual weeks. For simplicity, we extended the linear growth trend observed in early postnatal measurements to the perinatal period. Additionally, we introduced several constraints to the symbolic regression algorithm: no growth occurs at the initial stage ($g_r(\text{GA} = 29) = 1$), and the derivatives at the connection stage were required to be smooth ($g'_r(\text{GA} = 40) \in C^0$).

We constructed all computational models using publicly available 4D infant cortical surface atlases.^{60,61} These atlases also served as the base model for measuring quantities such as curvature, GI, and SulcDepth, providing a baseline for quantitative comparison. In our previous study, the cortical surface was parcellated into 18 distinct regions based on the NMF method.³⁰ For this study, however, the primary objective was to compare the effectiveness of the regional growth model with classic growth theories. Therefore, we selected five representative regions that exhibit significant distinctions in surface area

and cortical thickness. These regions are slow-growing region 1, medium-growing regions 4, 9, and 11, and fast-growing Region 16, corresponding to the cortical areas of “Dorsal Precentral, Paracentral and Posterior Cingulate”, “Lateral Precentral and Postcentral”, “Supra Marginal”, “Caudal Middle Frontal”, and “Dorsal Prefrontal”, respectively. As illustrated in Fig. 5, the development of the surface area and cortical thickness from 29 postmenstrual weeks to 2 years of age exhibits significant differences among these regions, with the p -values all smaller than 0.001. Here, we conducted paired Student's t -test to compute the significance of these differences, with the null hypothesis being no difference in surface area or cortical thickness between the two regions.

3. Results

In this study, we first employed a symbolic regression algorithm to identify suitable growth models using longitudinal surface area and cortical thickness data from developing human brains, with ages ranging from 29 postmenstrual weeks to 2 years of age. These growth models were then integrated into ABAQUS *via* a user-defined subroutine to simulate mechanical folding evolutions caused by differential growth in double-layer structures. The dynamic folding patterns were analyzed and compared with anatomically realistic brain imaging models both in qualitative and quantitative manner. Subsequently, we substituted the regional growth model with classical growth models commonly used in previous studies, such as isotropic growth and purely tangential growth. We evaluated and compared the effectiveness of each model in modeling cortical folding development using quantitative metrics introduced in Section 2.4. We further investigated whether the accuracy of modeling is more influenced by the value of growth ratio or the trajectory of the growth model.



Finally, we built a multi-regional model incorporating three regions with distinct growth patterns predicted by symbolic regression.

3.1. Regional growth models identified from symbolic regression

We investigated suitable tangential growth models through symbolic regression using the in-plane growth ratio data, normalized by the initial growth ratio as shown in eqn (6). This normalization ensures that the growth ratio starts from an initial value of 1. Moreover, given that the cortical surface area continues to increase beyond 24 postnatal months,^{62,63} the derivative of the growth ratio at this point must remain positive ($g'_t > 0$). These constraints were incorporated into the symbolic regression algorithm by customizing the loss function to impose a significant penalty for any violation. Fig. 6 illustrates the tangential growth model identified for each representative region using symbolic regression. All models exhibit satisfactory accuracies, with R^2 values exceeding 0.85, effectively capturing the growth patterns characterized by an initial linear increasing phase transitioning into a gradually steady phase, forming an upper-sigmoid shape. Intriguingly, among the predefined operators such as “exp”, “tanh”, “ln”, and polynomials, as introduced in Section 2.2, the symbolic regression consistently selected a combination of tanh and polynomials.

Moreover, a consistent model format, $g_t(t) = (a \times t + b) \times \tanh(c \times t) + 1$, where a , b , and c are constants, was discovered for almost all regions except for region 1. This finding suggests a promising model choice for characterizing brain in-plane growth patterns.

For out-of-plane growth models, we applied distinct conditions to constrain growth behaviors. In addition to the initial value constraint ($g_r|_{t=0} = 1$), we incorporated an additional constraint to ensure the derivative is smooth at the connection stage ($g'_r|_{GA=40} \approx g'_r|_{t=0.1} \in C^0$) due to the use of linear extrapolation to address data scarcity during the perinatal period. Moreover, as the cortical thickness development tends to be stabilized within the first two years after birth,^{64,65} the derivative at the phase end must be zero ($g'_r|_{t=1} = 0$). Fig. 7 shows the radial growth models identified for each representative region through symbolic regression. Despite significant variance, the predicted growth model can still decipher the mathematical relationship from the provided data points. All growth models exhibit an inverted U-shape, peaking during the intermediate period. This phenomenon is consistent with imaging observations by Gilmore *et al.*⁶⁶ For growth in the out-of-plane directions, symbolic regression exclusively selected the polynomial operators to construct the base models. The general model format can be summarized as a combination of up to fourth order polynomials: $g_r(t) = a \times t^4 + b \times t^3 + c \times t^2 + d \times t + e$,

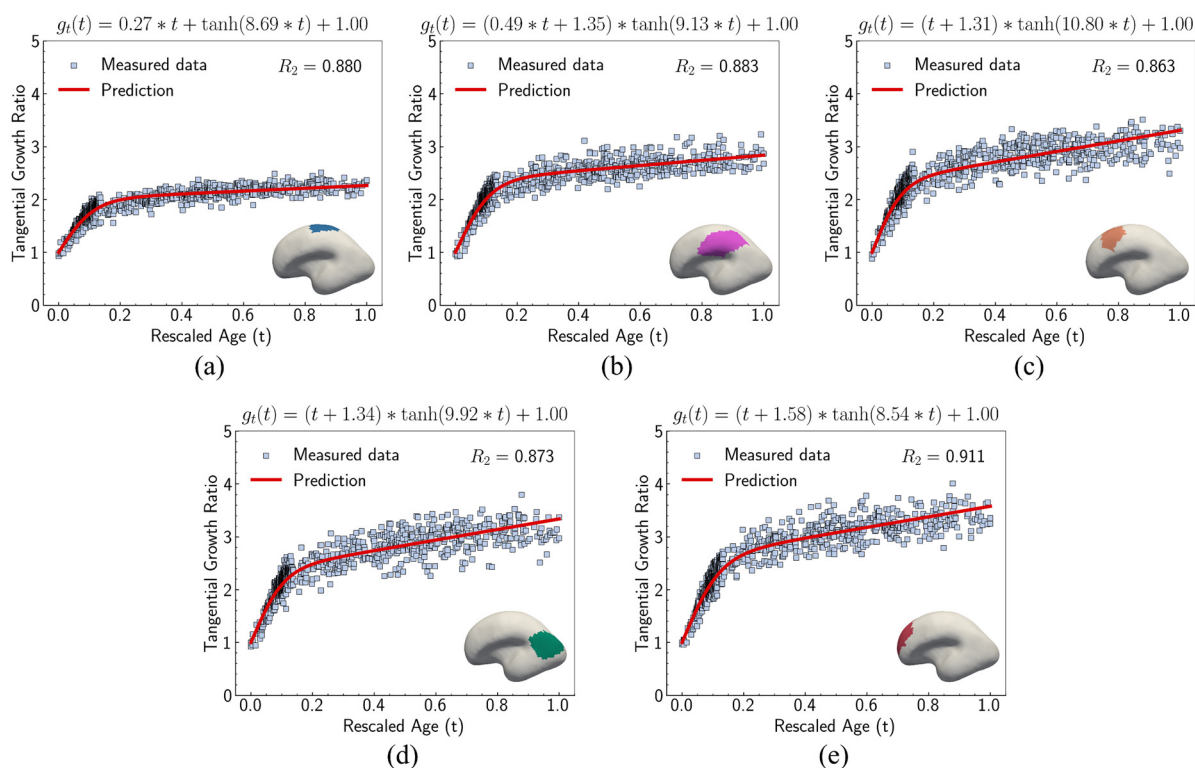


Fig. 6 Tangential growth models discovered for five selected regions. In-plane growth predicted by symbolic regression algorithm for (a) region 1, (b) region 4, (c) region 11, (d) region 9, and (e) region 16, respectively. R^2 indicates the goodness of fit. Mathematical forms of growth models are presented at the top of each figure. Insets show spatial locations of regions on the cortical surface. Dots denote MRI scans collected from 29 postmenstrual weeks to 2 years of age.



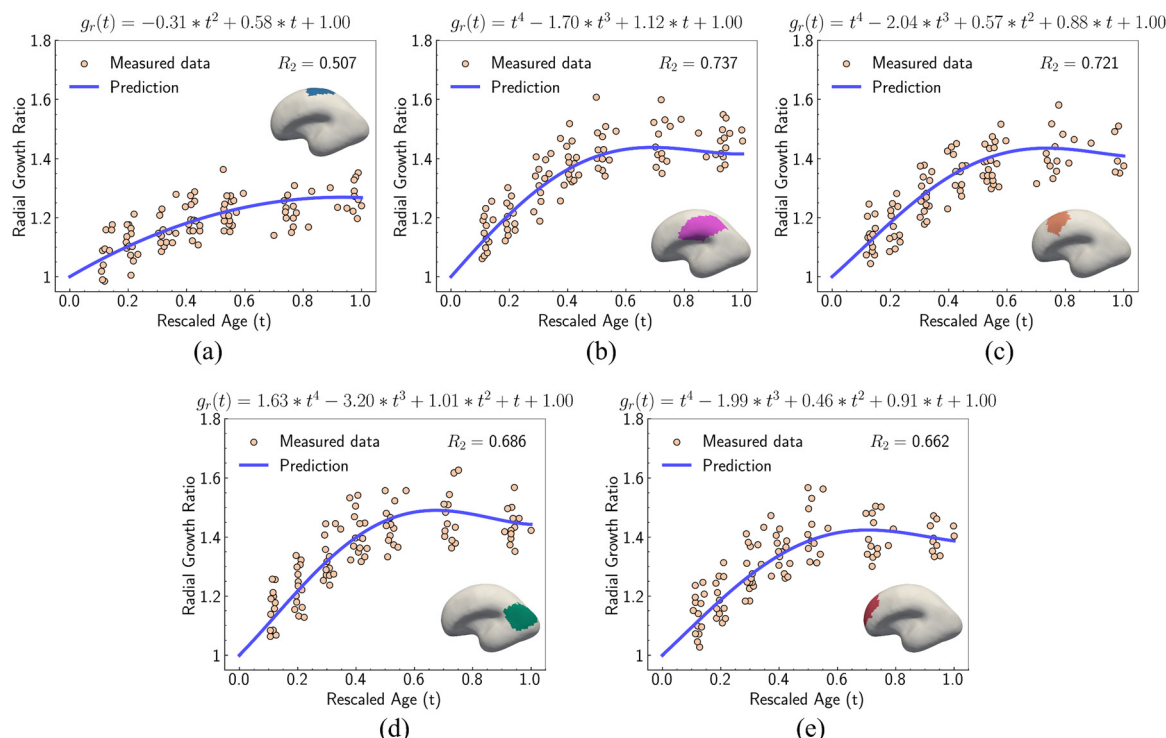


Fig. 7 Radial growth models discovered for five selected regions. Growth of thickness predicted by symbolic regression algorithm for (a) region 1, (b) region 4, (c) region 11, (d) region 9, and (e) region 16, respectively. R^2 indicates the goodness of fit. Mathematical forms of growth models are presented at the top of each figure. Insets show spatial locations of regions on the cortical surface. Dots denote MRI scans collected from 29 postmenstrual weeks to 2 years of age.

where a, b, c, d, e are constants. This result suggests a potential model choice for characterizing brain out-of-plane growth patterns. Though the accuracy was degraded by data variance, symbolic regression still proves robust in identifying suitable regional growth models for the human brain cortex, as validated in our recent paper.³⁹ These predicted models will be integrated into ABAQUS *via* user-defined subroutines to simulate brain folding evolution.

3.2. Regional growth models accurately simulate folding evolutions patterns

Fig. 8 illustrates the evolution of cortical folding on simulated brain surface, incorporating the regional growth model. The folding patterns at six key simulation moments are displayed and rendered with the displacement magnitude. As displayed, the non-uniform curvature inherent on initial brain surface (29 postmenstrual weeks) was sufficient to initiate bulking at around 0.2 s and generate intrinsic folding modes, which evolved to form distinct gyri and sulci over time. The resulting folding patterns of regional models are visually different. Specifically, regions 4, 11, and 16 exhibited denser folds compared to regions 1 and 9, indicating shorter folding wavelengths in these regions. This disparity is primarily attributed to differences in initial geometry and potentially the effects of growth. Additionally, for each model, areas with higher initial altitude (characterized in the out-of-plane direction) tend to experience greater displacements, as depicted in dark red

areas. Since we applied consistent growth ratios both in tangential and radial directions across each model, the non-uniform displacements likely stem from the in-plane compressions caused by tangential growth under boundary confinements. This suggests that tangential growth has a more significant influence on folding evolutions than radial growth, which is also evident in the larger values observed in tangential growth models compared to radial growth model in Fig. 6 and 7.

The nonuniform distribution of displacement remains prominent at the interface between the gray matter and white matter. As shown in Fig. 9a and b, greater displacements occur in areas with higher initial altitude. Additionally, the simulated folding surfaces exhibit characteristic pattern of cusped sulci and smooth gyri, consistent with the morphology observed in real human brain structures.⁶⁷ To validate our simulation results, we focused on the folding patterns at the interface where the initial locations approximately cover the extracted brain regions, as illustrated in Fig. 9c. We then compared these patterns with those of a 24-postnatal-month brain, as highlighted in Fig. 9d.⁶¹ As shown, our simulated brain folding closely replicates the realistic brain pattern, particularly in regions 1 and 4. For the other three regional models, certain characteristic morphologies are comparable to those in real brains, such as the ladder-shaped concaves in region 16 and the sharp ridge along with a narrow valley in region 9. To quantitatively compare these folding patterns, we further computed the curvature (MC), sulcal depth (SulcDepth), and gyrification



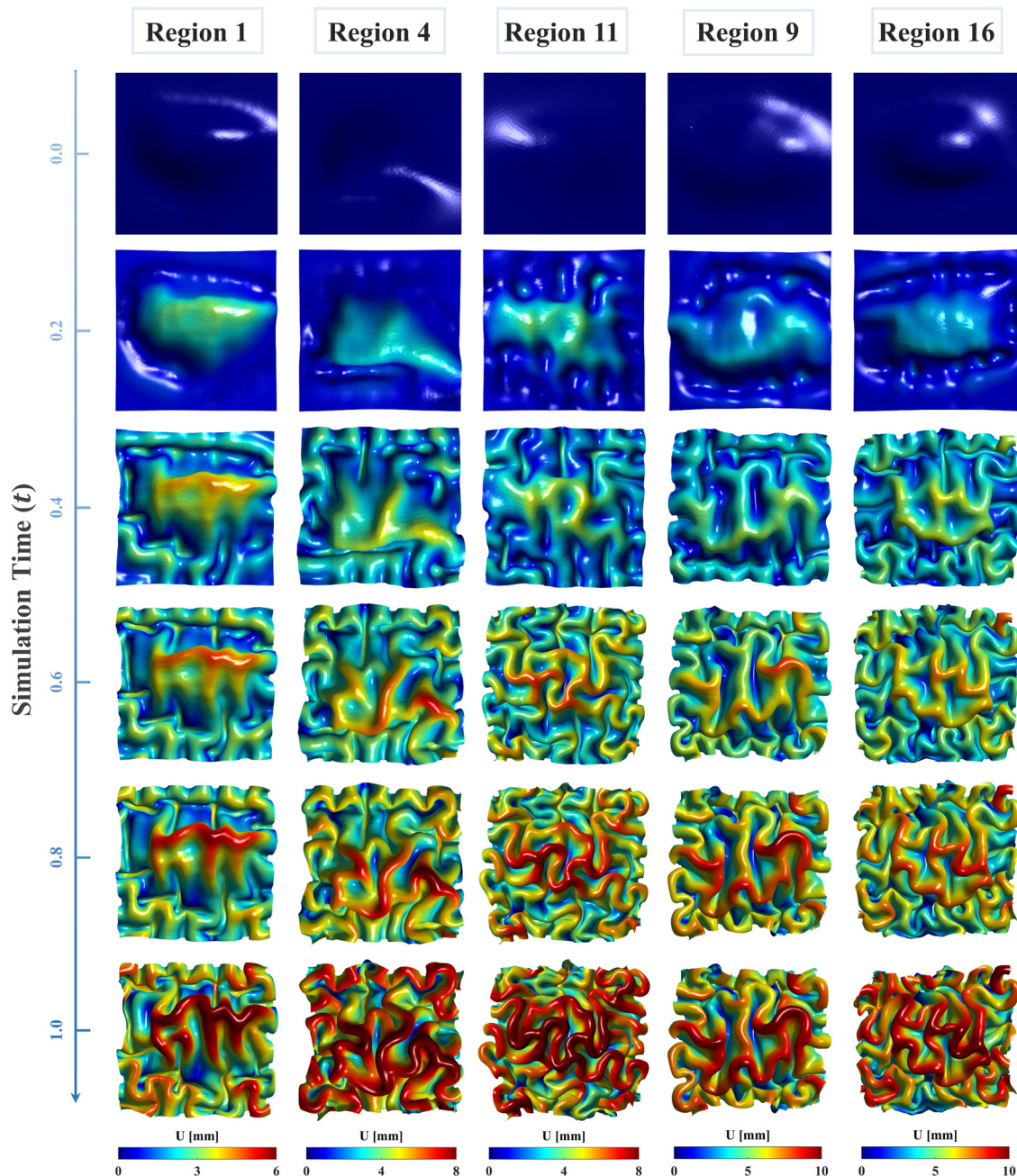


Fig. 8 Longitudinal brain developing patterns for five regions. Cortical folding patterns of six moments ranging from $t = 0$ to $t = 1.0$ were recorded and rendered with displacement magnitude.

index (GI) using the methods described in Section 2.4. The distributions of MC and SulcDepth are shown in the contour plot of Fig. 10. As depicted, positive curvatures tend to appear on the gyri, while negative on the sulci. This distinction suggests that mean curvature can serve as a metric to deliberately extract gyri and sulci for further investigations.^{45,50,68}

Fig. 11 shows the comparisons of those quantitative metrics measured in the simulated brain (circular dots and line) with those from a realistic human brain (square dots), with the dots representing the mean values of each metric averaged across the entire ROI surface. As shown in the figure, most of

simulation results align well with the realistic brain data in both trend and values. The simulation trajectories exhibit sigmoid-like shapes, starting with flat initial phases, then increasing dramatically after the onset of geometry instabilities (around 0.2 s, as seen in Fig. 8), and finally stabilizing once the folding patterns mature. This trend aligns with the identified regional growth models depicted in Fig. 6 and 7. Among the three metrics, mean curvature was poorly predicted compared to the other two metrics, as reflected in the insignificantly correlated trends and poor alignments, represented by low R^2 values equal to 0. It is important to note that the initial



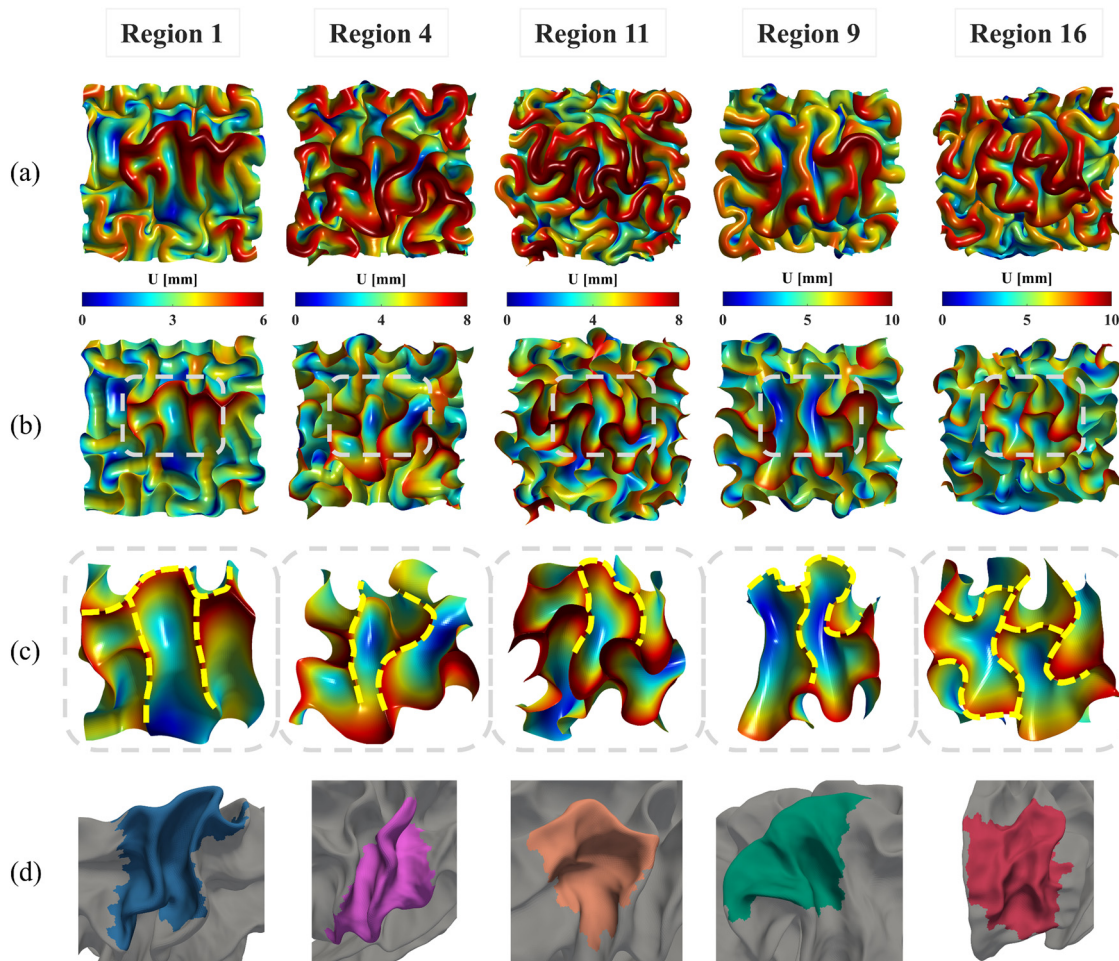


Fig. 9 Simulated folding patterns vs. realistic brain images. (a) Cortical folding patterns of five regions displayed on the cortical surface and rendered with displacement magnitude at $t = 1$ s; (b) cortical folding patterns of five regions displayed on the gray/white matter interface; (c) zoom-in representation of the highlighted regions associated with yellow dashed lines indicate the similar characteristic morphologies as observed in real human brain images (d).

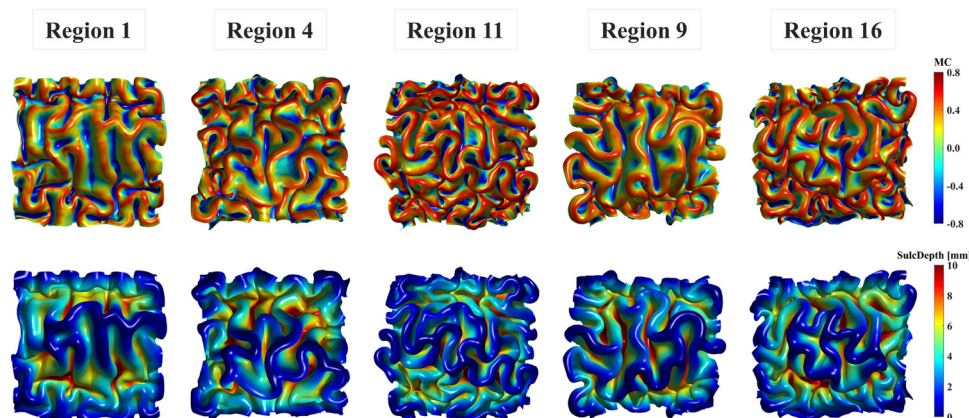


Fig. 10 Mean curvature and sulcal depth. Distribution of dimensionless mean curvature and sulcal depth of the five selected regions at the final state.

curvatures in the simulation are not zero because the simulation models were constructed based on brain regional surfaces, which already display curvatures at the initial state (29 post-menstrual weeks). The discrepancy between the two trajectories

can potentially be explained by the following factors: (1) absence of external constraints: the simulation does not incorporate the influence of external constraints from the brain skull, meninges, or cerebrospinal fluid. While these factors



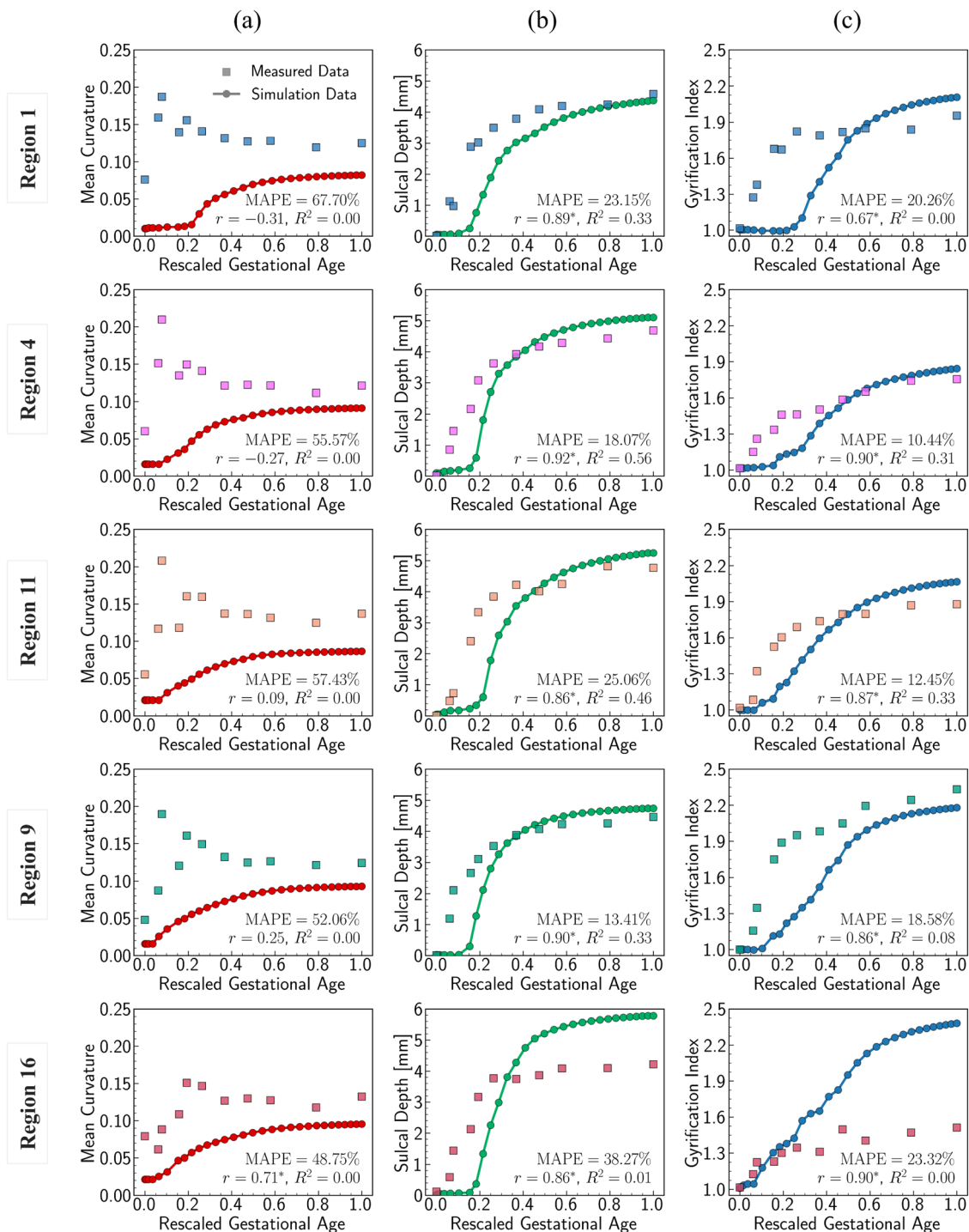


Fig. 11 Quantitative comparison of mean curvature, sulcal depth, and gyrification index among five regions. Mean curvature (a), sulcal depth (b), and gyrification index (c) of each moment. Ages of the measured data range from 29 postmenstrual weeks to 24 months after birth, and the gestational ages are rescaled into the range of [0, 1] to facilitate data comparison. Mean curvature is determined by averaging absolute dimensionless mean curvature at each point. The accuracy of the simulation results is evaluated using R^2 , with prediction error quantified by MAPE. The correlation between the simulation results and real brain data is analyzed using the Pearson coefficient (r), with statistically significant correlations ($p < 0.05$) marked by an asterisk (*). Squared dots represent data measured from human brain imaging data, while circular dots with lines represent data measured from simulated brain.

do not initiate cortical folding, they play an important role in regulating the brain's local geometry. For instance, contact with the skull and meninges may flatten the gyrus upon contact, resulting in wider and smoother gyral peaks.²⁰ (2) Inaccurate

sampling of the region of interest (ROI): for simplicity in postprocessing, we defined the ROI as the smallest rectangular area enclosing the extracted brain regions. This approach inevitably introduces redundant areas, thereby reducing the



precision of quantitative metrics, such as mean curvature, compared to measurements taken directly on the exact brain regions. The sulcal depth and gyrification index correlate well across all regions except for region 16, where a significant deviation occurs after $t = 0.4$ s. This discrepancy mainly arises from the geometric model being constructed based on a single region. Since region 16 exhibits the highest growth ratio among all selected models, it undergoes the largest deformations during simulation, which, under boundary confinement, would generate the most convoluted folding structures, resulting in the highest GI and sulcal depth. However, in real brain, the compression caused by rapid expansion of region 16 would be relaxed by the surrounding regions, which yields less complex folding patterns such as shallower sulci, as observed in Fig. 9. In summary, both qualitative and quantitative comparisons affirm the effectiveness of regional growth models in simulating the brain folding evolution process.

3.3. Regional growth models outperform classic unified growth models

To compare the effectiveness of regional growth models with classic models such as isotropic growth and purely tangential growth, we performed simulations on three regional models with varying growth speeds: the slow-growing region 1, medium-growing region 9, and fast-growing region 16. In both isotropic growth and tangential growth cases, the growth ratio was defined as a constant $g = \sqrt{8}$, which is commonly adopted in previous studies.^{25,53,67} The performances of different growth theories were compared in Fig. 12 (ESI,† Fig. S4), where folding patterns at $t = 1$ s were presented and rendered with displacement, mean curvature, and sulcal depth, respectively, to facilitate qualitative comparisons. Quantitative comparisons were also provided by averaging each metric across the entire ROI surface.

As shown in Fig. 12a–c, the regional growth model generates more sulcal and gyral folds compared to the isotropic growth model, especially evident in the intermediate areas corresponding to the initial brain surface. The increased number of folds indicates that regional growth model yields more convoluted folding patterns with shorter wavelengths. This phenomenon concurs with the findings of Budday and Steinmann,³² which demonstrates that variation in cortical growth modulates secondary instabilities, culminating in highly irregular folding patterns. However, the difference in wavelength between the regional growth model and the tangential growth model is not evident, even though the regional growth model tends to produce more cusped and deep sulci, as well as curled gyri. This suggests that tangential growth theory generates more convoluted patterns than isotropic growth, consistent with previous findings.^{69,70} Furthermore, this proves that tangential (in-plane) growth plays a more dominant role in affecting folding patterns than radial (out-of-plane) growth.

The quantitative comparison illustrated in Fig. 12d shows that the regional growth model provides more accurate folding measures compared to the other two growth theories. This is evidenced by the closer alignment of the resulting curvature, sulcal depth, and GI with real brain data, as reflected by lower

MAPE values. Additionally, the regional growth model initiates the folding earlier and reaches the stabilization stage sooner than the other two models. This suggests that the regional growth model can significantly reduce computational costs in simulating brain folding evolution by supplementing a stopping criterion based on stabilization.

3.4. Growth ratio values influence folding evolution more than growth trajectory

The regional growth model and purely tangential growth model exhibit similar primary folding patterns, with the regional growth model producing more cusped and deep sulci and curled gyri in secondary details. Both models adopt a similar format of tangential growth, but their values and growth trajectories differ. Noted, the effect of radial growth is ignored due to its minor impact on modulating folding. To figure out whether the value or growth trajectory dominates the folding patterns, we performed simulations on geometric model of region 9 using three growth models with the same resulting growth ratio but different growing trajectories. In addition to the regional growth model, we introduced two additional growth models following linear and Gompertz distributions, respectively, as shown in Fig. 13. The Gompertz distribution has been proven successful in modeling tissue growth,⁷¹ such as the growth mode of myelinated brain white matter⁷² and tumor growth with necrosis.⁷³ The Gompertz distribution is expressed as follows:

$$g_t(t) = a \times \exp(-\exp(-b \times (t - c))) + 1, \quad (11)$$

where a , b , and c are constants. Compared to the regional growth model, the Gompertz model exhibits a typical S shape, starting with an initial flat phase and quickly transitioning into a stabilization phase. In contrast, the linear model has a growth ratio that increases linearly.

Fig. 14 illustrates the qualitative comparison of folding patterns rendered with displacement, curvature, and sulcal depth, as well as the quantitative comparison by averaging each metric across the entire ROI surface. As seen in Fig. 14a–c, despite the difference in growing trajectories, the resulting folding demonstrates nearly identical patterns, especially in the linear and Gompertz models. This is also evident in quantitative measurements shown in Fig. 14d. Though the regional growth model initiates instabilities earlier due to its rapid growth within the first 0.2 s, the resulting measures of GI and curvature are almost the same across all models. Additionally, the regional growth model generates deeper sulci compared to the other two models. When compared to the findings in Section 3, it becomes apparent that the growth trajectories have a less significant impact on folding patterns than the magnitude of growth itself. This observation remains consistent with the findings of Wang *et al.*,⁴⁹ who report that the cortical growth mode has a limited influence on surface morphology complexity. Interestingly, this result appears to be an intrinsic feature of brain folding, as it persists across various simulation approaches. Our simulations were based on anatomically realistic brain models, incorporating a physical accurate growth pattern. In contrast, Wang *et al.*⁴⁹ employed



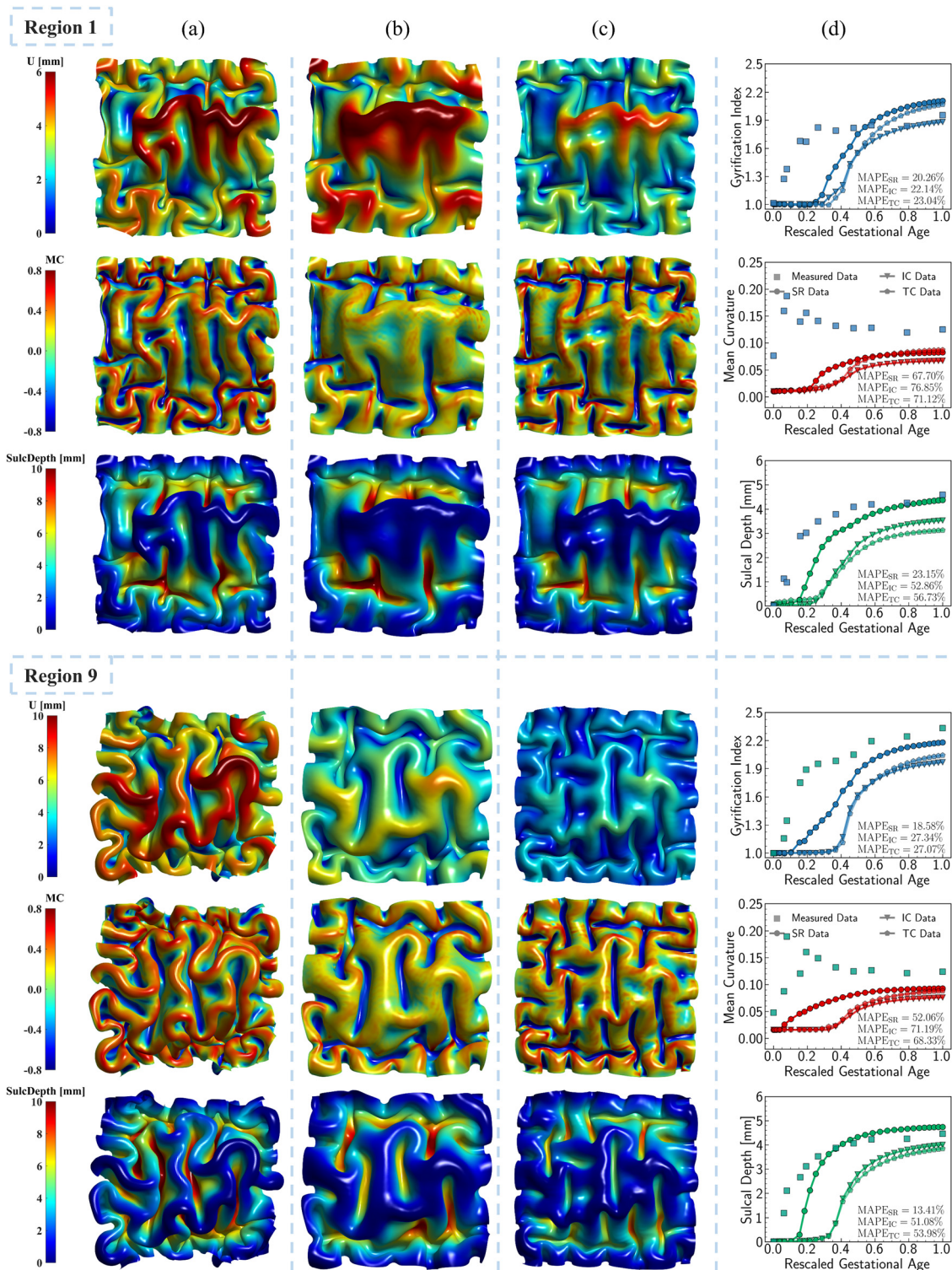


Fig. 12 Symbolic regression growth model vs. classic growth model. Impact of different growth models, including growth model predicted from symbolic regression (SR) algorithm (a), isotropic model with constant growth ratio (IC) model (b), and tangential model with constant growth ratio (TC) (c), on modeling cortical folding patterns. Distributions of displacement (U), mean curvature (MC), and sulcal depth ($SulcDepth$) of regions 1 and 9 are displayed with their quantitative measure present at the right (d). The simulation errors relative to real brain imaging data are quantified using MAPE. The growth ratio in both isotropic growth model is $g = \sqrt{8}$, while $g_t = \sqrt{8}$ in the tangential direction and no radial growth occurs for the pure tangential model, $g_r = 1$.



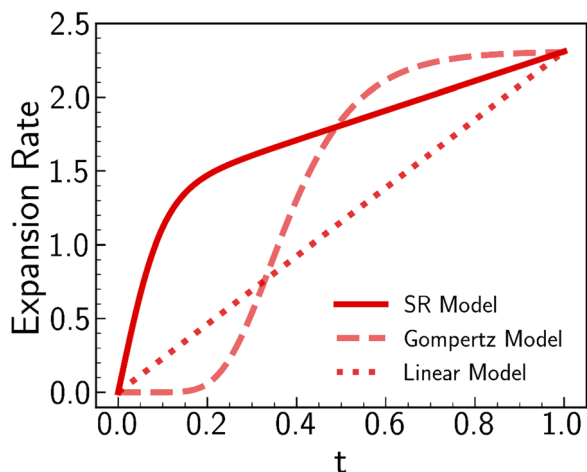


Fig. 13 Three distinct growth models for region 9. Initial and final values keep the same for all three models, respectively, while the developing trajectories are different.

idealized spherical models with manually defined growth patterns to represent brain structure. Despite these differences in methodology, the core finding—minor impact of growth trajectories on folding patterns—remains consistent.

3.5. Multi-region model provides more realistic folding results than single-regional model

All simulations discussed above were conducted on a single brain region, where we applied symmetric confinements to

their boundaries to prevent out-of-boundary growth. Despite our efforts to minimize the boundary effects by sufficiently extending the ROI boundaries, this condition may still be overly restrictive in generating compression, which is pivotal for triggering instabilities, resulting in unrealistic folding patterns as observed in region 16 (see Fig. 8 and 11). Additionally, in real brains, no region functions or deforms independently without affecting or being affected by surrounding regions. To address these limitations, we constructed a new computational model using the initial geometries of three adjacent regions: regions 1, 4, and 11. The constructed model is shown in Fig. 15a. After proper smoothing and extension, the model uniformly partitions a squared area into three parts, following an approximate proportion of 1 : 1 : 2 for regions 1, 11, and 4, respectively.

The simulated patterns are illustrated in Fig. 15b–d. Compared to the single-region model (see Fig. 8), the multi-region model generates more uniform folds across each region. For example, in the single-region modeling, the folding patterns are fairly condensed in the ROI of region 11 but sparse in region 1. Conversely, these patterns are uniformly distributed in the multi-region model, which is consistent with the anatomical morphology of the real human brain. However, certain characteristic features like the central sulcus were missed, which could be improved by incorporating more regions into the model construction process.

The quantitative comparison in Fig. 16 indicates that the multi-region model produces more realistic results compared to the single-region model. This is demonstrated by the closer

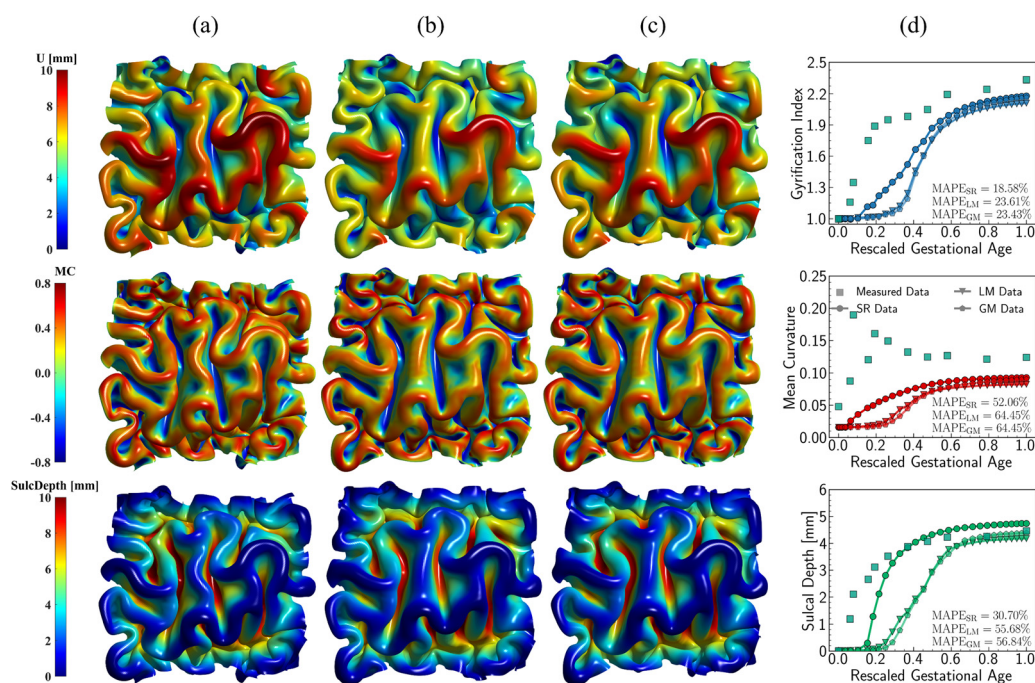


Fig. 14 Impact of growth trajectory on the folding patterns. Three growth models with distinct developing patterns for region 9, including symbolic regression model (SR) (a), linear growth model (LM) (b), and Gompertz growth model (GM) (c). Distribution of displacement (U), mean curvature (MC), and sulcal depth ($SulcDepth$) are present with their quantitative measure present at the right (d). The simulation errors relative to real brain imaging data are quantified using MAPE. In simulations, the above growth models are exclusively functioned in the tangential direction, while the growth in thickness remains consistent, and the radial growth model predicted from symbolic regression is employed, as shown in Fig. 6.



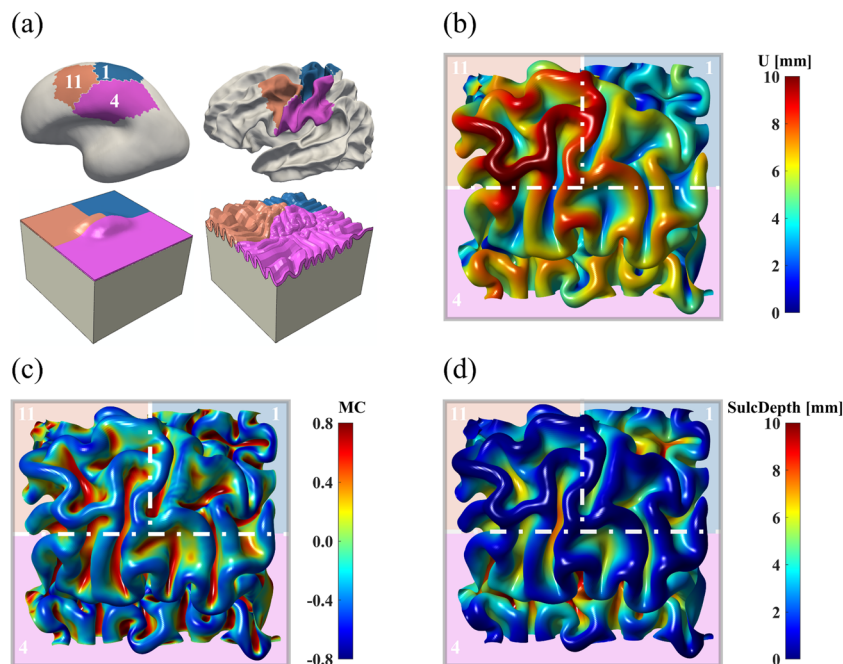


Fig. 15 Brain folding patterns of a multi-region model. (a) The computational model is based on regions 1, 4, and 11 of the human brain, with colored areas indicating distinct material properties, specifically growth behaviors. Cortical folding patterns in area of interest are displayed with displacement (b), mean curvature (c), and sulcal depth (d). Regions are delineated by white dotted lines and marked with distinct background colors.

alignment of quantitative measures for GI, mean curvature, and sulcal depth with real brain data, as evidenced by lower MAPE values and higher correlation coefficient in most cases. These results highlight the superiority of the multi-region model in simulating brain folding evolutions over the single-region model. Looking ahead, it is expected that a brain-wide model incorporating all 18 regions could yield more promising brain folding predictions, both in terms of qualitative patterns and quantitative measures.

4. Discussion

Computational simulations, such as finite element methods, have emerged as promising tools to elucidate the mechanical influence on cortical development, providing valuable insights into the underlying mechanisms^{22,67,74} and enabling predictive modeling of brain tissue behaviors under various conditions.^{75–77} Herein, we show that integrating simulations with regional growth models, derived through symbolic regression using brain developmental data, produces more realistic brain folding patterns. These region-based developing cortical patterns further exhibit accurate alignment with quantitative measures of real brain development. Our findings challenge the conventional views that isotropic, unified growth is sufficient to mimic cortical folding evolutions with modeling results more consistent with the imaging analyses.^{21,23} Instead, our results suggest that heterogeneous growth patterns play an indispensable role in modulating brain cortical development, as suggested in the works of Holland *et al.*⁷⁸ and Budday and Steinmann.³²

The longitudinal brain data encompass surface area expansion measured from 29 postmenstrual weeks to 2 years of age³⁰ and cortical thickness recorded within the first two years after birth.³¹ We employed symbolic regression to derive explicit mathematical expressions for growth models across various brain regions, spanning from slow-growing areas to fast-growing areas. The models identified exhibit consistent patterns in describing both in-plane and out-of-plane growth for nearly all regions. This uniformity in regional growth reflects a characteristic growth mode of neural tissues⁷⁹ and has also been observed in tumor development.^{71,80} Symbolic regression was chosen over conventional methods, such as multiple regression or neural networks, due to its ability to predict interpretable models from data without requiring prior knowledge of the model structure, which must be predefined in multiple regression.^{81,82} Moreover, symbolic regression explores an infinite functional space to identify candidate models, even with constraints applied to expedite the search process, highlighting its superior performance over other machine-learning-based methods,^{83,84} which are restricted to a limited functional space. The explicit mathematical expression of the growth model greatly facilitates the implementation of these models in ABAQUS simulations using user-defined subroutines. Though directly incorporating data into simulation can ensure closer alignments with the provided data, this approach is highly sensitive to noises, particular outliers, and may fail in scenarios requiring further operations, such as differentiation or integration.

Uniform growth theories such as isotropic growth have been extensively employed to model brain development.^{21–25} While



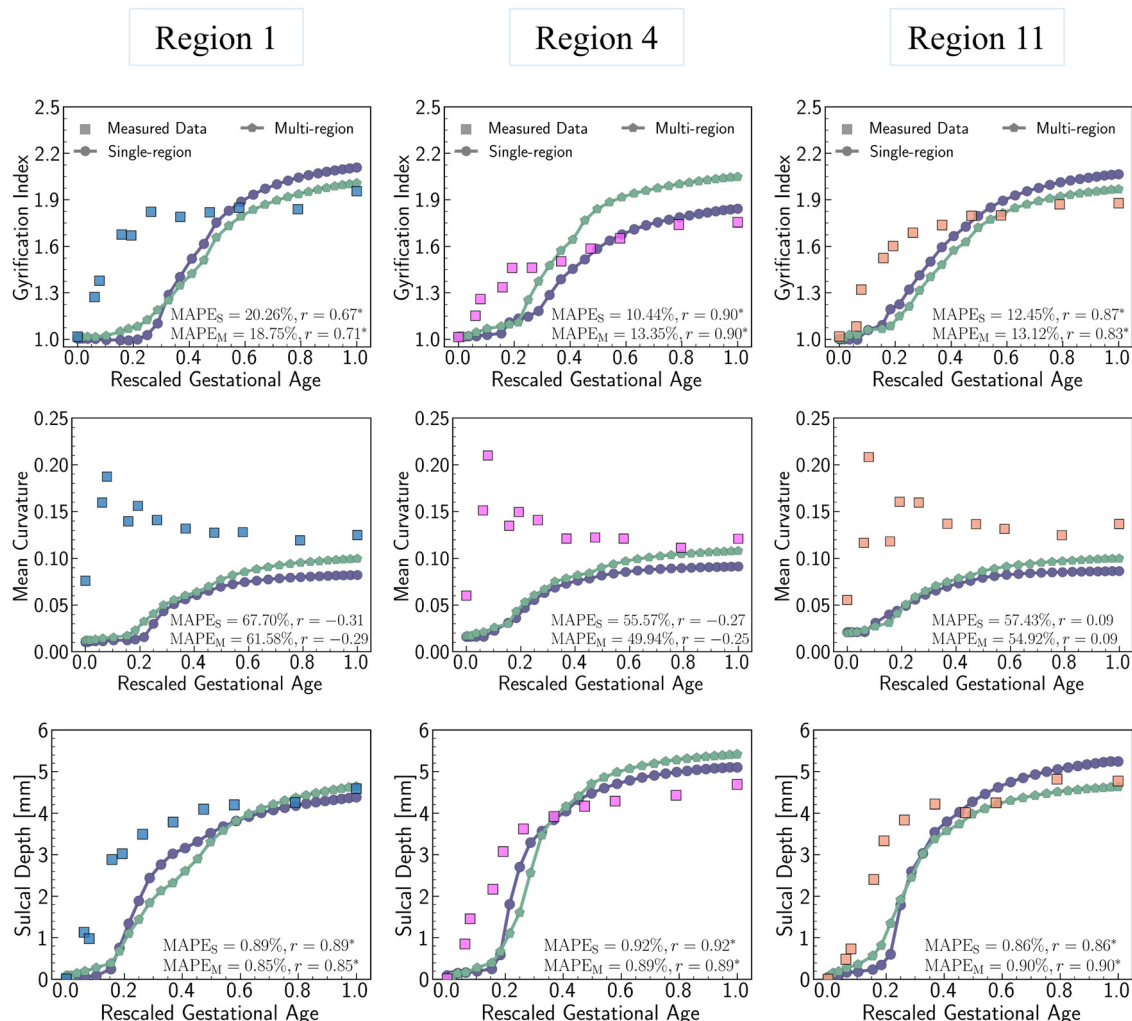


Fig. 16 Multi-region model vs. single-region model. Comparison on the trajectory of gyrification index, mean curvature, and sulcal depth between the multi-region model and single-region model. Mean curvature is determined by averaging absolute dimensionless mean curvature at each point. The simulation errors relative to real brain imaging data are quantified using MAPE, with MAPE_S and MAPE_M denoting errors for the single-region model and multi-region model, respectively. The correlation between simulation results and real brain data is assessed using the Pearson coefficient (r), with statistically significant correlations ($p < 0.05$) marked by an asterisk (*). Squared dots represent data measured from human brain imaging data, while circular dots and pentagonal dots with lines represent data measured from simulation based on single-region model and multi-region model, respectively.

these methods have significantly advanced our understanding of cortical folding mechanisms, their effectiveness degrades when comparing the resulting cortical folding patterns with imaging observations, particularly in brain-wide simulations.^{26,27} To address this limitation, additional physiological effects such as axonal fibers^{23,78} and cells migrations^{85,86} have been incorporated to more accurately simulate brain folding anatomy. However, these additions inevitably complicate modeling endeavors. Herein, we found that exclusively implementing regional growth into simulations can produce more realistic cortical folding patterns that closely match quantitative measurements of the real brain. This outcome aligns with the principle of Occam's Razor: the fewer deviations from real conditions, the better the simulation's alignment with the observed data. This finding underscores the superiority of heterogeneous growth over homogeneous growth in brain modeling.²⁰ As presented in

Section 3.3, even purely tangential growth can yield more complex folding patterns, similar to those observed in regional growth scenarios, compared to isotropic growth. Introducing such heterogeneity decreases the system's stability by triggering earlier bulking and more convoluted folding patterns. In addition to the regional growth, heterogeneity also exists in other brain properties like stiffness,³² surface curvature,⁸⁷ and cortical thickness.⁸⁸ These factors collectively contribute to the complex evolution of brain folding patterns.

While the brain surface bulking is predominantly driven by intrinsic properties such as initial surface curvature, relative stiffness or thickness ratio between gray matter and white matter, the growth does influence the cortical folding, especially in terms of cortical depth and folding complexity, as revealed by quantitative measures. Moreover, the value of growth ratio plays a more crucial role in modulating folding



evolutions compared to the growth trajectory. Intuitively, the external forces, such as the thermal expansion introduced here, do not dictate the bulking modes but rather influence the timing of bulking initialization and the extent of bulking development. However, brain folding is a dynamic process, with numerous complex factors interacting in a coupled manner. Heterogeneous growth can alter these intrinsic features. For example, differential growth in the in-plane and out-of-plane directions can change the thickness ratios between the gray matter and white matter. Similarly, axonal maturation, influenced by cell migration, can affect tissues stiffness.⁸⁹ These features evolve dynamically, introducing variability into the cortical folding process.

Conducting simulations on a single brain region allows for comparative analysis of distinct growth theories and their effects on brain folding. However, this approach may lead to unrealistic folding patterns due to the artificially imposed boundary conditions. Our findings indicate that a multi-region computational model, which considers three adjacent regions simultaneously, offers a more reliable result by producing more uniformly distributed folding patterns. In the future, a brain-wide model encompassing all 18 parcellated regions is expected to yield more realistic folding predictions, by integrating regional growth models derived through symbolic regression. Moreover, our study can be improved by addressing the following issues: first, we assumed uniform cortical thickness across each brain region. Incorporating anatomically accurate cortical thickness, accounting for both gray and white matter in the model construction process, would provide a more convincing geometric model. Second, the tangential growth within the cortical layer was assumed to be uniform. In reality, this growth varies spatially, as evident in differential growth within the six-layered cortex.⁹⁰ Future studies should consider adopting a spatially dependent growth profile, as proposed by Tallinen *et al.*⁸ Third, in current study, the tangential and radial growth models were characterized based on different datasets, as shown in Fig. 6 and 7. Future studies that integrate surface area and cortical thickness measurements from the same dataset would significantly enhance the integrity and rigor of the predicted growth model. Last but not least, the brain tissue in our model was treated as an incompressible hyperelastic material described by the neo-Hookean strain energy function. Incorporating a regional hyperelastic model with a degree of compressibility, characterized through symbolic regression, could account for the heterogeneity in stiffness and further enhance the reliability of our simulation results.

5. Conclusion

In this study, we investigated the impact of regional growth on the computational simulation of brain cortical folding. Using symbolic regression, we derived explicit mathematical growth models using the longitudinal data of cortical surface area expansion and cortical thickness measured on real human brains. These regressed models were then integrated into

ABAQUS to simulate the evolution process of cortical folding. The resulting folding patterns were recorded and quantified using several mechanical descriptors, including mean curvature, sulcal depth, and gyrification index. Our findings demonstrate that regional growth models can generate complex folding patterns that closely resemble the anatomical structure of the brain in both quantitative and qualitative aspects. These models outperformed conventional uniform growth models such as isotropic growth and purely tangential growth in accurately modeling brain folding developments. We observed that the magnitude of growth, rather than its trajectory, plays a more predominant role in shaping folding patterns. Moreover, multi-region modeling produces more uniform folding patterns similar to imaging observations compared to single-region models. Our results underscore the importance of implementing regional growth in brain folding simulations. This approach holds promise for advancing early diagnostics of cortical malformations like pachygyria, lissencephaly, and polymicrogyria and improves treatment for neurodevelopmental disorders such as autism and epilepsy.

Author contributions

JH: methodology, software, validation, investigation, writing – original draft; ZW: data curation, writing – review and editing; XC: formal analysis, writing – review and editing; LW: validation, writing – review and editing; DZ: validation, writing – review and editing; TL: validation, writing – review and editing; GL: conceptualization, validation, writing – review and editing; XW: conceptualization, validation, supervision, funding acquisition, writing – original draft, writing – review and editing.

Data availability

The original contributions presented in the study are included in the article/ESI†. Further inquiries can be directed to the corresponding authors. The dHCP dataset is publicly available at the Developing Human Connectome Project repository: <https://www.developingconnectome.org>.⁹¹ The BCP dataset is publicly available in NIMH Data Archive: https://nda.nih.gov/edit_collection.html?id=2848.⁹²

Conflicts of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

JH and XW acknowledge the support from National Science Foundation (IIS-2011369) and National Institutes of Health (1R01NS135574-01). GL is supported in part by National Institutes of Health (MH123202, ES033518, AG075582, NS128534, and NS135574). This work also utilizes approaches developed



by an NIH grant (1U01MH110274) and the efforts of the UNC/UMN Baby Connectome Project Consortium. Some of the data was provided by the developing Human Connectome Project, KCL-Imperial-Oxford Consortium funded by the European Research Council under the European Union Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. [319456]. We are grateful to the families who generously supported this research.

References

- G. Lohmann, D. Y. von Cramon and A. C. F. Colchester, *Cereb. Cortex*, 2007, **18**, 1415–1420.
- J. E. Schmitt, A. Raznahan, S. Liu and M. C. Neale, *Cereb. Cortex*, 2021, **31**, 702–715.
- J. Zeidan, E. Fombonne, J. Scorsah, A. Ibrahim, M. S. Durkin, S. Saxena, A. Yusuf, A. Shih and M. Elsabbagh, *Autism Res.*, 2022, **15**, 778–790.
- E. Beghi, *Neuroepidemiology*, 2020, **54**, 185–191.
- R. A. McCutcheon, T. R. Marques and O. D. Howes, *JAMA Psychiatry*, 2020, **77**, 201–210.
- L. Subramanian, M. E. Calcagnotto and M. F. Paredes, *Front. Cell. Neurosci.*, 2020, **13**, 576.
- D. C. Van Essen, *Semin. Cell Dev. Biol.*, 2023, **140**, 90–104.
- T. Tallinen, J. Y. Chung, J. S. Biggins and L. Mahadevan, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 12667–12672.
- E. Takahashi, R. D. Folkerth, A. M. Galaburda and P. E. Grant, *Cereb. Cortex*, 2011, **22**, 455–464.
- R. M. Fame and M. K. Lehtinen, *Dev. Cell*, 2020, **52**, 261–275.
- L. F. Franchini, *Front. Cell Dev. Biol.*, 2021, **9**, 591017.
- P. V. Bayly, L. A. Taber and C. D. Kroenke, *J. Mech. Behav. Biomed. Mater.*, 2014, **29**, 568–581.
- C. D. Kroenke and P. V. Bayly, *J. Neurosci.*, 2018, **38**, 767–775.
- D. C. V. Essen, *Nature*, 1997, **385**, 313–318.
- J. Nie, L. Guo, K. Li, Y. Wang, G. Chen, L. Li, H. Chen, F. Deng, X. Jiang, T. Zhang, L. Huang, C. Faraco, D. Zhang, C. Guo, P.-T. Yap, X. Hu, G. Li, J. Lv, Y. Yuan, D. Zhu, J. Han, D. Sabatinelli, Q. Zhao, L. S. Miller, B. Xu, P. Shen, S. Platt, D. Shen, X. Hu and T. Liu, *Cereb. Cortex*, 2011, **22**, 2831–2839.
- G. Xu, A. K. Knutsen, K. Dikranian, C. D. Kroenke, P. V. Bayly and L. A. Taber, *J. Biomech. Eng.*, 2010, **132**, 071013.
- L. Ronan, N. Voets, C. Rua, A. Alexander-Bloch, M. Hough, C. Mackay, T. J. Crow, A. James, J. N. Giedd and P. C. Fletcher, *Cereb. Cortex*, 2014, **24**, 2219–2228.
- M. R. Rosenzweig, *Dev. Neuropsychol.*, 2003, **24**, 523–540.
- D. P. Richman, R. M. Stewart, J. Hutchinson and V. S. Caviness Jr, *Science*, 1975, **189**, 18–21.
- M. Darayi, M. E. Hoffman, J. Sayut, S. Wang, N. Demirci, J. Consolini and M. A. Holland, *J. Biomech.*, 2022, **139**, 110851.
- P. V. Bayly, R. J. Okamoto, G. Xu, Y. Shi and L. A. Taber, *Phys. Biol.*, 2013, **10**, 016005.
- S. Budday, P. Steinmann and E. Kuhl, *J. Mech. Phys. Solids*, 2014, **72**, 75–94.
- P. Chavoshnejad, X. Li, S. Zhang, W. Dai, L. Vasung, T. Liu, T. Zhang, X. Wang and M. J. Razavi, *Brain Multiphys.*, 2021, **2**, 100029.
- M. Holland, S. Budday, A. Goriely and E. Kuhl, *Phys. Rev. Lett.*, 2018, **121**, 228002.
- M. J. Razavi, T. Liu and X. Wang, *Cerebral Cortex Commun.*, 2021, **2**, tgab044.
- J. Nie, L. Guo, G. Li, C. Faraco, L. Stephen Miller and T. Liu, *J. Theor. Biol.*, 2010, **264**, 467–478.
- S. N. Verner and K. Garikipati, *Extreme Mech. Lett.*, 2018, **18**, 58–69.
- K. E. Garcia, X. Wang, S. E. Santiago, S. Bakshi, A. P. Barnes and C. D. Kroenke, *Cereb. Cortex*, 2024, **34**, bhae172.
- C. D. Kroenke, E. N. Taber, L. A. Leigland, A. K. Knutsen and P. V. Bayly, *Cereb. Cortex*, 2009, **19**, 2916–2929.
- Y. Huang, Z. Wu, F. Wang, D. Hu, T. Li, L. Guo, L. Wang, W. Lin and G. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2121748119.
- F. Wang, C. Lian, Z. Wu, H. Zhang, T. Li, Y. Meng, L. Wang, W. Lin, D. Shen and G. Li, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 15855–15860.
- S. Budday and P. Steinmann, *Int. J. Solids Struct.*, 2018, **132**, 31–41.
- S. Wang, N. Demirci and M. A. Holland, *Biomech. Model. Mechanobiol.*, 2021, **20**, 555–567.
- T. Zhang, M. J. Razavi, X. Li, H. B. Chen, T. M. Liu and X. Q. Wang, *Sci. Rep.*, 2016, **6**, 37272.
- E. K. Rodriguez, A. Hoger and A. D. McCulloch, *J. Biomech.*, 1994, **27**, 455–467.
- G. A. Holzapfel, *Nonlinear solid mechanics: a continuum approach for engineering science*, Kluwer Academic Publishers, Dordrecht, 2002.
- D. Angelis, F. Sofos and T. E. Karakasidis, *Arch. Comput. Methods Eng.*, 2023, **30**, 3845–3865.
- B. Bahmani and W. Sun, *Int. J. Numer. Meth. Eng.*, 2024, **125**, e7473.
- J. Hou, X. Chen, T. Wu, E. Kuhl and X. Wang, *Acta Biomater.*, 2024, **188**, 276–296.
- Z. Zhang, Z. Zou, E. Kuhl and G. E. Karniadakis, *Comput. Methods Appl. Mech. Eng.*, 2024, **419**, 116647.
- W. Ben Chaabene and M. L. Nehdi, *Constr. Build. Mater.*, 2021, **280**, 122523.
- M. Cranmer, *arXiv*, 2023, preprint, arXiv:2305.01582, DOI: [10.48550/arXiv.2305.01582](https://doi.org/10.48550/arXiv.2305.01582).
- J. Dubois, M. Alison, S. J. Counsell, L. Hertz-Pannier, P. S. Hüppi and M. J. Benders, *J. Magn. Reson. Imaging*, 2021, **53**, 1318–1343.
- G. Abaqus, *Dassault Systemes Simulia Corporation*, Providence, RI, USA, 2011, vol. 3.
- P. Chavoshnejad, L. Chen, X. Yu, J. Hou, N. Filla, D. Zhu, T. Liu, G. Li, M. J. Razavi and X. Wang, *Cereb. Cortex*, 2023, **33**, 9354–9366.
- M. Jalil Razavi, M. Reeves and X. Wang, *Comput. Methods Biomech. Biomed. Eng.*, 2017, **20**, 1212–1222.
- J. Weickenmeier, R. de Rooij, S. Budday, T. C. Ovaert and E. Kuhl, *J. Mech. Behav. Biomed. Mater.*, 2017, **76**, 119–124.



- 48 Y. Cao, Y. Jiang, B. Li and X. Feng, *Acta Mech. Solida Sin.*, 2012, **25**, 483–492.
- 49 X. Wang, J. Lefèvre, A. Bohi, M. A. Harrach, M. Dinomais and F. Rousseau, *Sci. Rep.*, 2021, **11**, 7686.
- 50 S. Zhang, P. Chavoshnejad, X. Li, L. Guo, X. Jiang, J. Han, L. Wang, G. Li, X. Wang, T. Liu, M. J. Razavi, S. Zhang and T. Zhang, *Hum. Brain Mapp.*, 2022, **43**, 4540–4555.
- 51 D. Duan, S. Xia, I. Rekik, Z. Wu, L. Wang, W. Lin, J. H. Gilmore, D. Shen and G. Li, *Hum. Brain Mapp.*, 2020, **41**, 1985–2003.
- 52 M. Meyer, M. Desbrun, P. Schröder and A. H. Barr, *Discrete Differential-Geometry Operators for Triangulated 2-Manifolds*, Berlin, Heidelberg, 2003.
- 53 R. Balouchzadeh, P. V. Bayly and K. E. Garcia, *Brain Multiphys.*, 2023, **4**, 100065.
- 54 H. J. Yun, K. Im, J.-J. Yang, U. Yoon and J.-M. Lee, *PLoS One*, 2013, **8**, e55977.
- 55 A. Makropoulos, E. C. Robinson, A. Schuh, R. Wright, S. Fitzgibbon, J. Bozek, S. J. Counsell, J. Steinweg, K. Vecchiato and J. Passerat-Palmbach, *NeuroImage*, 2018, **173**, 88–112.
- 56 E. J. Hughes, T. Winchman, F. Padormo, R. Teixeira, J. Wurie, M. Sharma, M. Fox, J. Hutter, L. Cordero-Grande and A. N. Price, *Magn. Reson. Med.*, 2017, **78**, 794–804.
- 57 J. De Asis-Cruz, J.-H. Kim, D. Krishnamurthy, C. Lopez, K. Kapse, N. Andescavage, G. Vezina and C. Limperopoulos, *Dev. Cognit. Neurosci.*, 2023, **63**, 101282.
- 58 L. Zubiaurre-Elorza, S. Soria-Pastor, C. Junque, R. Sala-Llloch, D. Segarra, N. Bargallo and A. Macaya, *PLoS One*, 2012, **7**, e42148.
- 59 N. Demirci and M. A. Holland, *Cereb. Cortex*, 2024, **34**, bhad462.
- 60 G. Li, L. Wang, F. Shi, J. H. Gilmore, W. Lin and D. Shen, *Med. Image Anal.*, 2015, **25**, 22–36.
- 61 Z. Wu, L. Wang, W. Lin, J. H. Gilmore, G. Li and D. Shen, *Hum. Brain Mapp.*, 2019, **40**, 3860–3880.
- 62 A. E. Lyall, F. Shi, X. Geng, S. Woolson, G. Li, L. Wang, R. M. Hamer, D. Shen and J. H. Gilmore, *Cereb. Cortex*, 2014, **25**, 2204–2212.
- 63 H. Kalantar-Hormozi, R. Patel, A. Dai, J. Ziolkowski, H.-M. Dong, A. Holmes, A. Raznahan, G. A. Devenyi and M. M. Chakravarty, *NeuroImage*, 2023, **268**, 119885.
- 64 H. Huang, R. Xue, J. Zhang, T. Ren, L. J. Richards, P. Yarowsky, M. I. Miller and S. Mori, *J. Neurosci.*, 2009, **29**, 4263–4273.
- 65 M. Ouyang, T. Jeon, A. Sotiras, Q. Peng, V. Mishra, C. Halovanic, M. Chen, L. Chalak, N. Rollins, T. P. L. Roberts, C. Davatzikos and H. Huang, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 4681–4688.
- 66 J. H. Gilmore, R. C. Knickmeyer and W. Gao, *Nat. Rev. Neurosci.*, 2018, **19**, 123–137.
- 67 T. Tallinen, J. Y. Chung, F. Rousseau, N. Girard, J. Lefèvre and L. Mahadevan, *Nat. Phys.*, 2016, **12**, 588–593.
- 68 T. Zhang, H. Chen, M. J. Razavi, Y. Li, F. Ge, L. Guo, X. Wang and T. Liu, *Hum. Brain Mapp.*, 2018, **39**, 4134–4149.
- 69 P. Chavoshnejad, L. Vallejo, S. Zhang, Y. Guo, W. Dai, T. Zhang and M. J. Razavi, *Sci. Rep.*, 2023, **13**, 13177.
- 70 K. E. Garcia, C. D. Kroenke and P. V. Bayly, *Philos. Trans. R. Soc., B*, 2018, **373**, 30249772.
- 71 M. Peterson, B. C. Warf and S. J. Schiff, *J. Neurosurg. Pediatr.*, 2018, **21**, 478–485.
- 72 N. Sadeghi, M. Prastawa, P. T. Fletcher, J. Wolff, J. H. Gilmore and G. Gerig, *NeuroImage*, 2013, **68**, 236–247.
- 73 C. Vaghi, A. Rodallec, R. Fanciullino, J. Ciccolini, J. P. Mochel, M. Matri, C. Poignard, J. M. Ebos and S. Benzekry, *PLoS Comput. Biol.*, 2020, **16**, e1007178.
- 74 M. Jalil Razavi, T. Zhang, T. Liu and X. Wang, *Sci. Rep.*, 2015, **5**, 14477.
- 75 M. Alenyà, X. Wang, J. Lefèvre, G. Auzias, B. Fouquet, E. Eixarch, F. Rousseau and O. Camara, *Brain Multiphys.*, 2022, **3**, 100045.
- 76 K. E. Garcia, X. Wang and C. D. Kroenke, *Nat. Commun.*, 2021, **12**, 6681.
- 77 M. A. Holland, S. Budday, G. Li, D. Shen, A. Goriely and E. Kuhl, *Eur. Phys. J.: Spec. Top.*, 2020, **229**, 2757–2778.
- 78 M. A. Holland, K. E. Miller and E. Kuhl, *Ann. Biomed. Eng.*, 2015, **43**, 1640–1653.
- 79 E. M. W. Billig, W. P. O'Meara, E. M. Riley and F. E. McKenzie, *Malar. J.*, 2012, **11**, 64.
- 80 K. R. Swanson, C. Bridge, J. Murray and E. C. Alvord Jr, *J. Neurol. Sci.*, 2003, **216**, 1–10.
- 81 N. Filla, J. Hou, T. Liu, S. Budday and X. Wang, *J. Mech. Behav. Biomed. Mater.*, 2024, **150**, 106271.
- 82 J. Hou, N. Filla, X. Chen, M. J. Razavi, T. Liu and X. Wang, *arXiv*, 2023, preprint, arXiv:2310.10762, DOI: [10.48550/arXiv.2310.10762](https://doi.org/10.48550/arXiv.2310.10762).
- 83 M. Flaschel, S. Kumar and L. De Lorenzis, *Comput. Methods Appl. Mech. Eng.*, 2023, **405**, 115867.
- 84 K. Linka, S. R. S. Pierre and E. Kuhl, *Acta Biomater.*, 2023, **160**, 134–151.
- 85 R. d Rooij and E. Kuhl, *J. Mech. Phys. Solids*, 2018, **112**, 563–576.
- 86 M. S. Zarzor, S. Kaessmair, P. Steinmann, I. Blümcke and S. Budday, *Brain Multiphys.*, 2021, **2**, 100025.
- 87 S. Budday, P. Steinmann, A. Goriely and E. Kuhl, *Extreme Mech. Lett.*, 2015, **4**, 193–198.
- 88 N. Demirci and M. A. Holland, *Hum. Brain Mapp.*, 2022, **43**, 2064–2084.
- 89 J. Guo, G. Bertalan, D. Meierhofer, C. Klein, S. Schreyer, B. Steiner, S. Wang, R. V. da Silva, C. Infante-Duarte and S. Koch, *Acta Biomater.*, 2019, **99**, 433–442.
- 90 S. Wang, K. Saito, H. Kawasaki and M. A. Holland, *PLoS Comput. Biol.*, 2022, **18**, e1010190.
- 91 D. H. C. Project, Accessed 2024.
- 92 B. R. Howell, M. A. Styner, W. Gao, P.-T. Yap, L. Wang, K. Baluyot, E. Yacoub, G. Chen, T. Potts and A. Salzwedel, *NeuroImage*, 2019, **185**, 891–905.

