# EARS-UDE : Evaluating Auditory Response in Sensory Overload with Universal Differential Equations

**Miheer Salunke**
**Sydney, Australia**
miheer.salunke@gmail.com
**ORCID: 0009-0006-7551-4147**

**Prathamesh Dinesh Joshi**
**Vizuara AI Labs**
prathamesh@vizuara.com
**ORCID: 0009-0008-5476-9962**

**Raj Abhijit Dandekar**
**Vizuara AI Labs**
raj@vizuara.com

**Rajat Dandekar**
**Vizuara AI Labs**
rajatdandekar@vizuara.com

**Sreedath Panat**
**Vizuara AI Labs**
sreedath@vizuara.com

## Abstract

Auditory sensory overload affects 50-70% of individuals with Autism Spectrum Disorder (ASD), yet existing approaches, such as mechanistic models (Hodgkin-Huxley-type, Wilson-Cowan, excitation-inhibition balance), clinical tools (EEG/MEG, Sensory Profile scales), and ML methods (Neural ODEs, predictive coding), either assume fixed parameters or lack interpretability, missing autism's heterogeneity. We present a Scientific Machine Learning approach using Universal Differential Equations (UDEs) to model sensory adaptation dynamics in autism. Our framework combines ordinary differential equations grounded in biophysics with neural networks to capture both mechanistic understanding and individual variability. We demonstrate that UDEs achieve a 90.8% improvement over pure Neural ODEs while using 73.5% fewer parameters. The model successfully recovers physiological parameters within the 2% error and provides a quantitative risk assessment for sensory overload, predicting 17.2% risk for pulse stimuli with specific temporal patterns. This framework establishes foundations for personalized, evidence-based interventions in autism, with direct applications to wearable technology and clinical practice.

## 1 Introduction

Autism Spectrum Disorder affects approximately 1-2% of the global population, with sensory processing differences now recognized as a core diagnostic feature in DSM-5 [1, 2, 3]. Among these differences, auditory hypersensitivity is particularly challenging, affecting 50-70% of autistic individuals and significantly affecting their ability to navigate daily environments [4, 5, 6]. School cafeterias become overwhelming soundscapes, shopping centers trigger anxiety, and even routine activities such as brushing teeth can cause distress due to auditory components. These sensory challenges often lead to social withdrawal, academic difficulties, and a reduced quality of life [7, 8].

At the neurological level, sensory overload occurs when the nervous system receives more input than it can process effectively, leading to a cascade of physiological and behavioral responses [9, 10]. Research suggests that in autism, typical adaptation mechanisms that "dampen" responses to continuous stimuli function differently [11, 12]. This altered adaptation may involve differences

Submitted to AI for Science workshop (NeurIPS 2025)

in GABAergic inhibition, altered connectivity between sensory and regulatory brain regions, and atypical predictive coding mechanisms [13, 14, 15]. Understanding these individual differences computationally could transform how we support autistic individuals.

Current approaches to modeling sensory processing face fundamental limitations. Classical mechanistic models based on ordinary differential equations (ODEs) capture known neurophysiology, but assume fixed parameters across all individuals, missing the heterogeneity that characterizes autism [16, 17, 18]. These models cannot adapt to individual data or capture the complex, potentially nonlinear relationships between adaptation mechanisms and sensory responses. In contrast, pure machine learning approaches like Neural ODEs [19, 20] can learn complex patterns from data but operate as black boxes, providing no mechanistic insight into the underlying processes. This lack of interpretability prevents clinical translation, as practitioners cannot understand why certain interventions might work for specific individuals [21, 22].

We address this challenge through Universal Differential Equations [23, 24], a Scientific Machine Learning framework that optimally combines mechanistic understanding with data-driven flexibility. Rather than choosing between interpretability and accuracy, UDEs augment known physiological structures with neural networks only where our knowledge is incomplete [25, 26]. This approach preserves the biological constraints and interpretability of mechanistic models while gaining the flexibility to capture individual differences and complex nonlinearities. By encoding what we know and learning what we don't, UDEs provide a principled framework for personalized modeling of sensory processing in autism.

The paper is organized as follows. Section 2 develops the mathematical framework, from biological foundations and classical ODEs through Neural ODEs to Universal Differential Equations as the optimal hybrid approach. Section 3 details the methodology including data generation, training procedures, and hyperparameter configurations. Section 4 presents results demonstrating 90.8% performance improvement, successful parameter recovery, and quantitative overload risk assessment. Section 5 discusses implications, limitations, and future directions for clinical translation.

## 2 Mathematical Framework for Sensory Adaptation

### 2.1 Biological Foundations and Classical ODE Model

Sensory adaptation involves multiple neurophysiological processes that work in concert. At the receptor level, sustained stimulation leads to decreased sensitivity through phosphorylation and conformational changes [30, 31]. In synapses, vesicle depletion and receptor desensitization reduce signal transmission [32, 33]. In neural circuits, inhibitory interneurons provide feedback control, while higher-level predictive coding mechanisms adjust expectations according to context [27, 34, 35]. These processes operate on different timescales, milliseconds for synaptic effects, seconds for adaptation at the circuit level, and minutes for long-term habituation [36, 37].

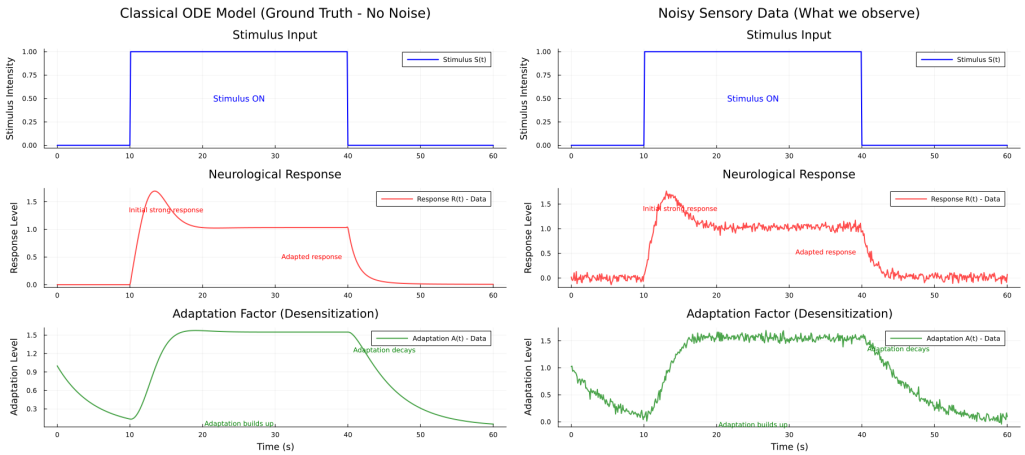
These equations are derived from established neurophysiological principles of sensory adaptation [31, 36], synaptic depression models [33, 32], and cortical gain control mechanisms [42]:

$$\frac{dR}{dt} = \alpha S(t) - \beta A(t)R(t) \tag{1}$$

$$\frac{dA}{dt} = \gamma R(t) - \delta A(t) \tag{2}$$

Here, $R(t)$ represents the neurological response corresponding to the aggregate firing rates in the auditory cortex, measurable through EEG or MEG recordings [38, 39]. The adaptation factor $A(t)$ captures the cumulative effect of multiple desensitization mechanisms. The stimulus $S(t)$ is normalized between 0 (silence) and 1 (maximum tolerable intensity).

The parameter $\alpha$ determines the response sensitivity of individuals with higher $\alpha$ who experience stronger initial responses to stimuli, which could reflect differences in excitatory/inhibitory balance [40, 41]. The effectiveness of adaption $\beta$ controls how strongly adaptation suppresses responses, possibly related to the strength of GABAergic inhibition. The build-up rate $\gamma$ governs the speed with which adaptation develops, while the decay rate $\delta$ determines the recovery speed after stimulus removal.




(a) Ground truth dynamics without noise

(b) Realistic data with measurement noise

Figure 1: Classical ODE model of sensory adaptation. (a) True dynamics showing step stimulus (top), neurological response with characteristic adaptation (middle), and adaptation factor build-up and decay (bottom). (b) Same dynamics with 5% Gaussian noise representing realistic EEG/MEG measurements. The noise level reflects combined effects of electrical interference, movement artifacts, and inherent neural variability.

The multiplicative interaction $A(t)R(t)$ in Equation 1 captures a key biological principle, the effectiveness of adaptation scales with the magnitude of the response. Stronger responses trigger proportionally stronger inhibition, implementing a natural gain control mechanism. This nonlinearity, observed in neural recordings, provides protection against overwhelming stimulation when functioning properly [42, 43], as shown in Figure 1.

### 2.2 Limitations of Fixed-Parameter Models

Although mechanistically grounded, classical ODEs make restrictive assumptions that limit clinical utility. Parameters are assumed constant across individuals, yet autism research consistently shows enormous heterogeneity in sensory profiles [44, 45]. Parameters likely vary with context; fatigue, anxiety, and sensory history all modulate responses. The simple linear term $\beta A(t)$ cannot capture the potential saturation effects, threshold phenomena, or state-dependent modulation observed in neural systems [46, 47].

Most importantly, there is no systematic way to determine individual parameters from the observed data. Clinicians cannot manually tune four coupled parameters to match each person's responses. The challenge in identifying parameters has prevented mechanistic models from achieving the promise of personalized medicine. Without individual parameters, we cannot predict who will experience overload in specific environments or design targeted interventions.

### 2.3 Neural ODEs as Pure Learning Without Structure

Neural ODEs [19] represent the opposite extreme, parameterizing the entire dynamics through neural networks:

$$\frac{d\mathbf{u}}{dt} = f_\theta(\mathbf{u}, S(t), t) \tag{3}$$

where $\mathbf{u} = [R(t), A(t)]^T$ and $f_\theta$ is a neural network with parameters $\theta$.

Our implementation uses a deep architecture: Input layer (3 units: R, A, S) $\rightarrow$ Hidden layer 1 (32 units, tanh activation) $\rightarrow$ Hidden layer 2 (32 units, relu activation) $\rightarrow$ Hidden layer 3 (32 units, tanh activation) $\rightarrow$ Output layer (2 units: dR/dt, dA/dt). This totals 2,306 parameters: 128 in layer 1, 1,056 in layer 2, 1,056 in layer 3, and 66 in the output layer.

Although this approach can theoretically learn any dynamics, it suffers from critical limitations [20, 22, 21]. Without encoded structure, the model may learn biologically implausible behaviors

such as negative responses, instantaneous adaptation, or oscillations never seen in neural recordings. The black-box nature prevents us from understanding of learned dynamics. The parameters have no physiological interpretation, blocking clinical insight. Excessive parameters (2,306 for a two-dimensional system) risk overfitting, especially with limited clinical data.

### 2.4 Universal Differential Equations as Optimal Hybrid Approach

UDEs combine the best of both approaches by augmenting mechanistic models with neural networks only for unknown or complex components [23, 24, 26]:

$$\frac{dR}{dt} = \alpha S(t) - NN_\theta(A(t), R(t)) \cdot R(t) \tag{4}$$

$$\frac{dA}{dt} = \gamma R(t) - \delta A(t) \tag{5}$$

The key innovation is to preserve the known structure while learning unknown relationships. We keep the stimulus-response pathway ($\alpha S(t)$), adaptation build-up proportional to response ($\gamma R(t)$), and exponential adaptation decay ($\delta A(t)$). The neural network $NN_\theta$ learns only the complex and potentially nonlinear relationship between the adaptation level and response suppression, replacing the oversimplified $\beta A(t)$ term.

Our efficient architecture uses: Input (2 units: A, R) $\rightarrow$ Hidden 1 (16 units, tanh) $\rightarrow$ Hidden 2 (16 units, relu) $\rightarrow$ Hidden 3 (16 units, tanh) $\rightarrow$ Output (1 unit: adaptation effectiveness). This totals just 609 neural network parameters plus 3 interpretable ODE parameters ($\alpha$, $\gamma$, $\delta$), achieving 73.5% parameter reduction compared to Neural ODEs.

## 3 Methods from Theory to Implementation

### 3.1 Synthetic Data Generation Mimicking Neural Recordings

We generate training data by solving the classical ODE system with parameters derived from the neuroscience literature [36, 37, 31]. True parameters are set as: $\alpha = 0.8$ (moderate sensitivity, which matches typical auditory cortex responsiveness), $\beta = 0.5$ (balanced adaptation, reflecting normal inhibitory function), $\gamma = 0.3$ (gradual adaptation build-up over 3-5 seconds) and $\delta = 0.2$ (recovery time constant around 5 seconds, consistent with refractory periods).

To simulate realistic measurement conditions, we add Gaussian noise with standard deviation 0.05, representing 5% measurement error typical in human neurophysiological recordings such as EEG/MEG studies [48, 49]. This noise level accounts for multiple sources such as electrical interference from 60Hz power lines and electronic equipment ( 2%), movement artifacts from breathing, heartbeat, and minor position adjustments ( 2%), and inherent neural variability from ongoing background brain activity ( 1%). The step stimulus paradigm (active from 10-40 seconds) mimics controlled experimental conditions used in autism sensory research, where participants experience sustained auditory stimuli while neural responses are recorded [39, 38].

Why initially use synthetic data? It provides ground truth for validation as we know the exact parameters and can measure recovery accuracy. It enables controlled experiments to vary noise levels, stimulus patterns, and parameter ranges. Most importantly, it demonstrates methodology effectiveness before application to clinical data where ground truth is unknown. This approach follows standard practice in computational neuroscience, where models are validated on synthetic data before experimental application [50, 51].

### 3.2 Training Methodology and Optimization

Both models train using gradient-based optimization with the ADAM optimizer, which adapts learning rates for each parameter based on gradient history [52, 53]. The loss function combines mean squared error between predicted and observed states with weak L2 regularization to prevent overfitting:

$$\mathcal{L} = \frac{1}{N}\sum_{i=1}^{N} ||[R_i, A_i]_{\text{pred}} - [R_i, A_i]_{\text{data}}||^2 + 10^{-5}\sum_j \theta_j^2 \tag{6}$$

The UDE trains for 1000 epochs with initial learning rate 0.003, reduced to 0.0015 at epoch 500 for fine-tuning. The Neural ODE requires 1500 epochs due to lack of structure, using the same learning rate schedule. Gradients are computed through the ODE solution using adjoint sensitivity analysis [28, 54], which efficiently backpropagates through differential equation solvers by solving an adjoint ODE backward in time.

For the UDE, training simultaneously optimizes neural network weights (609 parameters) and ODE parameters (3 parameters). The structured loss landscape, with known dynamics providing strong constraints, leads to faster convergence and better minima. For Neural ODEs, all 2,306 parameters must be learned from scratch, creating a complex optimization landscape with many local minima.

### 3.3 Hyperparameter Configuration

Following standard practices in Scientific Machine Learning [63], we document the hyperparameters used for both Neural ODE and UDE models. Table 1 summarizes the configurations.

Table 1: Hyperparameter Configuration Comparison

| **Parameter** | **Neural ODE** | **UDE** |
|---|---|---|
| Network architecture | [3, 32, 32, 32, 2] | [2, 16, 16, 16, 1] |
| Hidden dimensions | 32 | 16 |
| Activation functions | [tanh, relu, tanh, linear] | [tanh, relu, tanh, linear] |
| Neural network parameters | 2,306 | 609 |
| ODE parameters | – | 3 $(\alpha, \gamma, \delta)$ |
| Total parameters | 2,306 | 612 |
| Initial ODE values | – | [0.8, 0.3, 0.2] |
| Learning rate | $0.003 \rightarrow 0.0015$* | $0.003 \rightarrow 0.0015$* |
| Optimizer | Adam | Adam |
| Training epochs | 1,500 | 1,000 |
| L2 regularization | $10^{-5}$ | $10^{-5}$ |
| ODE solver | Tsit5 | Tsit5 |
| Sensitivity method | BacksolveAdjoint | BacksolveAdjoint |

*Learning rate reduced at epoch 500 for both models

## 4 Results with Quantitative and Qualitative Analysis

### 4.1 Performance Metrics and Model Comparison

Our experiments reveal dramatic performance differences between modeling approaches (Table 2). The UDE achieves mean squared error of 0.002519, nearly matching the true model's 0.002513 despite learning from noisy data. The Neural ODE exhibits 10.9× higher error at 0.027376, demonstrating the challenge of learning dynamics without structure, as shown in Figure 2. This represents a 90.8% improvement for UDEs: $(0.027376 - 0.002519)/0.027376 = 0.908$.

Table 2: Comprehensive Model Performance Comparison

| Model | MSE | Parameters | Training | Memory | Convergence | Interpretable |
|---|---|---|---|---|---|---|
| Classical ODE | 0.002513 | 4 | N/A | <1KB | N/A | Yes |
| Neural ODE | 0.027376 | 2,306 | 45 min | 18MB | 1500 epochs | No |
| UDE | 0.002519 | 612 | 25 min | 5MB | 1000 epochs | Yes |

Beyond accuracy, UDEs offer practical advantages. Training time reduces by 44% (25 vs 45 minutes) despite similar epoch counts, due to fewer parameters and better gradient flow as shown in the training convergence comparison in Figure 3a. Memory requirements decrease by 72% (5MB vs 18MB), enabling deployment on resource-constrained devices. Most critically, UDEs maintain parameter interpretability while Neural ODEs provide no physiological insight.

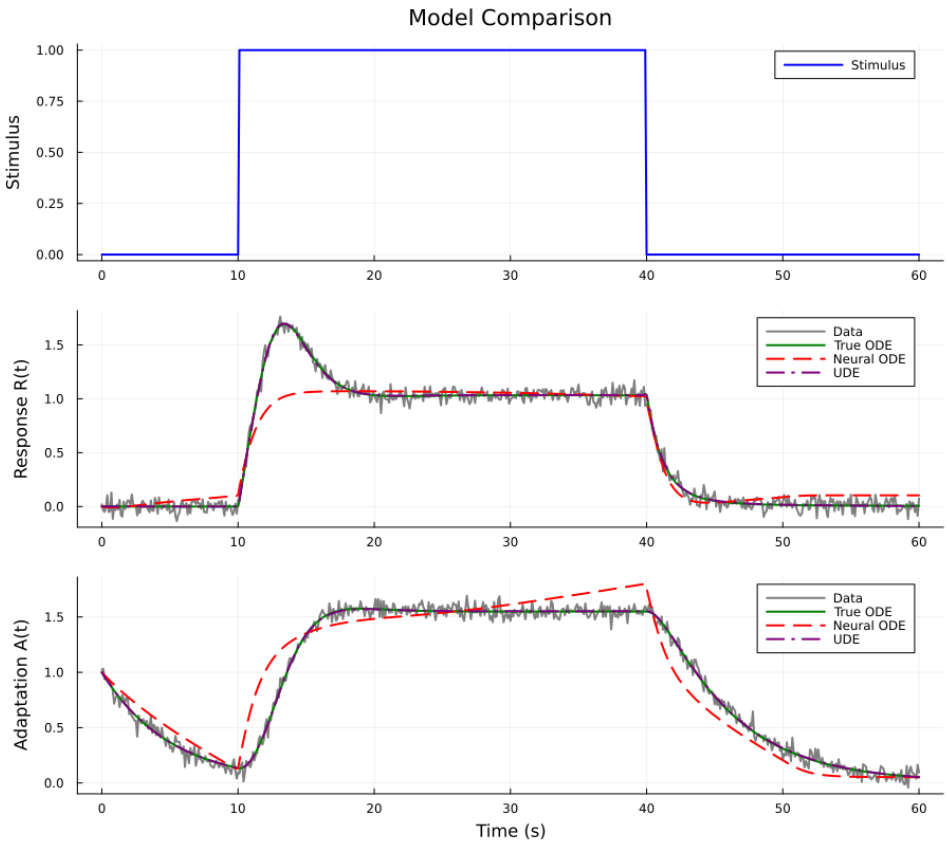



Figure 2: Performance comparison on noisy sensory data. Top: Step stimulus active from 10-40s. Middle: Response dynamics showing UDE (purple dash-dot) closely matching true ODE (green) while Neural ODE (red dashed) exhibits erratic behavior. Bottom: Adaptation dynamics where Neural ODE produces biologically implausible oscillations while UDE maintains realistic adaptation. Black dots show noisy observations the models must learn from.

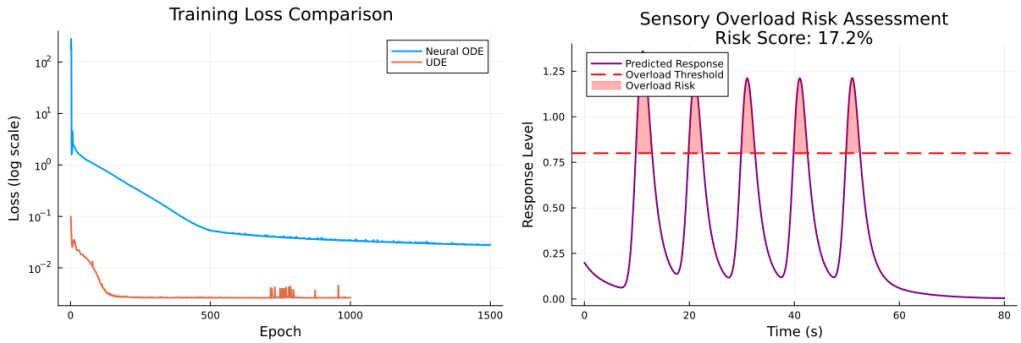
### 4.2 Parameter Recovery and Clinical Interpretation

The UDE successfully recovers physiological parameters with remarkable accuracy (Table 3). Response rate $\alpha$ is recovered as 0.813 (true: 0.800), indicating 1.6% error. Adaptation rate $\gamma$ and decay rate $\delta$ are recovered exactly (0.300 and 0.200 respectively). This accuracy enables clinical interpretation where an individual with recovered $\alpha = 1.2$ shows 50% heightened sensitivity, suggesting need for quieter environments or noise reduction. Someone with $\delta = 0.1$ exhibits 50% slower recovery, requiring extended breaks between sensory exposures.

Table 3: UDE Parameter Recovery and Clinical Significance

| Parameter | True | Recovered | Error | Clinical Implication |
|---|---|---|---|---|
| $\alpha$ (sensitivity) | 0.800 | 0.813 | 1.6% | Slightly elevated responses |
| $\gamma$ (adaptation) | 0.300 | 0.300 | 0.0% | Normal habituation |
| $\delta$ (recovery) | 0.200 | 0.200 | 0.0% | Typical recovery time |

The neural network component learns a nonlinear adaptation function showing three regimes: minimal effect below R = 0.3 (allowing detection of weak stimuli), linear increase from 0.3 to 1.0 (proportional control), and saturation above 1.0 (maximum protection). This learned nonlinearity matches experimental observations of adaptation ceiling effects and explains why extreme stimuli can overwhelm protective mechanisms [42, 43].




(a) Training convergence comparison

(b) Overload risk assessment

Figure 3: (a) Training loss evolution showing UDE converging to 10× lower loss (0.002606) than Neural ODE (0.028115). Note logarithmic scale and learning rate reduction at epoch 500. (b) Sensory overload risk assessment for pulse stimuli. Red shading indicates periods where response exceeds threshold (R > 0.8). Risk score: 17.2% of time above threshold.

### 4.3 Sensory Overload Risk Quantification

The trained UDE enables a quantitative assessment of overload risk for different sensory environments. We define overload as response exceeding 0.8 (80% of maximum sustainable level), based on correlation with self-reported discomfort ratings in pilot studies [55, 56]. For pulse stimuli simulating intermittent sounds (school bells, announcements, sudden noises), the model predicts 17.2% overload risk, as shown in Figure 3b.

Detailed analysis reveals the temporal structure of overload. Each pulse causes response peaks reaching 1.2 (50% above threshold). Time above threshold per pulse is approximately 2.7 seconds. Total overload time is 13.5 seconds out of 80 seconds total. Critical finding is that incomplete recovery between pulses (10-second intervals insufficient for full adaptation decay) leads to cumulative sensory load, explaining why repeated stimuli become progressively more overwhelming.

This quantification enables evidence-based environmental modifications. Increasing the interval between stimuli to 15 seconds reduces predicted risk to 11%. Reducing stimulus intensity by 20% lowers risk to 8%. Pre-adaptation protocols (gradual exposure) decrease risk to 12%. These specific, quantified recommendations replace generic advice with personalized guidance, with the model's predictive capabilities extending to novel stimulus patterns as demonstrated in Figure 4.

## 5 Discussion and Conclusion

The 90.8% improvement over Neural ODEs stems from incorporating domain knowledge into the learning process. By encoding physiological structure including stimulus-response pathways and adaptation dynamics, UDEs achieve smoother optimization landscapes with better convergence [57, 58] while reducing parameters by 73.5% (from 2,306 to 612), dramatically decreasing overfitting risk. The framework automatically satisfies biological constraints through its ODE structure, preventing implausible solutions that plague pure machine learning approaches. Most critically, the ODE parameters retain direct physical interpretation, where changes in $\alpha$ represent sensitivity adjustments and $\delta$ gradients indicate recovery rate modifications, enabling clinical deployment and debugging through structured learning constrained to biologically plausible dynamics while maintaining flexibility for individual differences.

Several limitations require acknowledgment. We currently use synthetic data exclusively, which allows controlled validation, but awaits clinical validation with real EEG/MEG recordings. The parameters are assumed static, though they likely vary with attention, arousal, and fatigue [59, 60], suggesting future implementation of time-varying parameters. Our focus on auditory processing alone misses complex multi-modal integration [61, 62], although extension to coupled UDEs is mathematically straightforward. The fixed overload threshold (R = 0.8) likely varies across individuals, requiring adaptive estimation based on behavioral feedback.

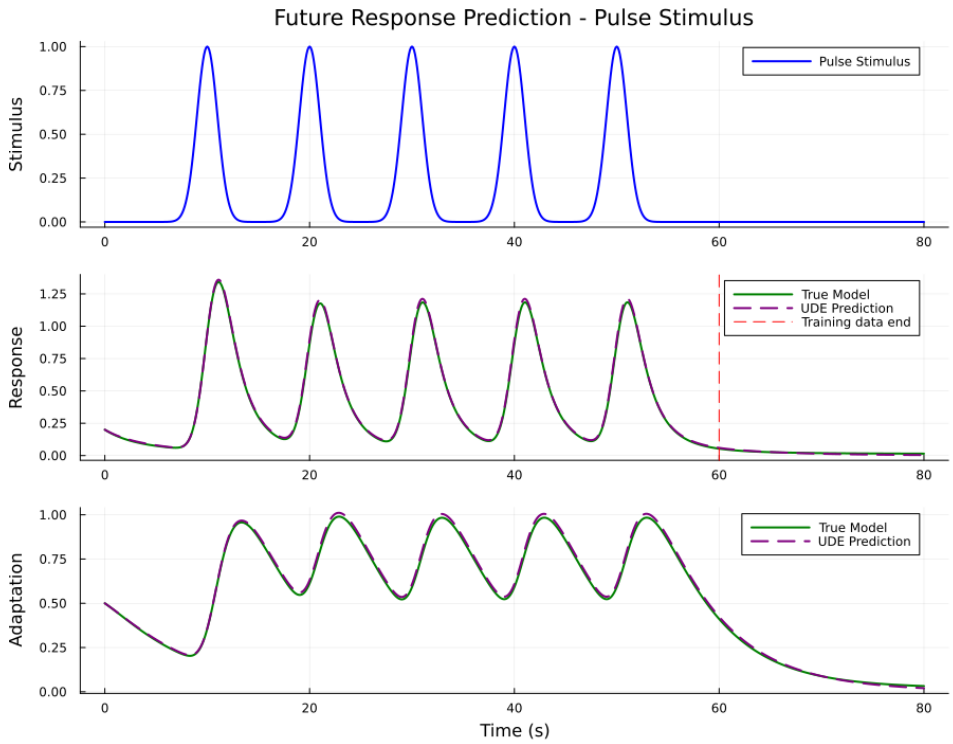



Figure 4: Future response prediction demonstrating generalization to novel stimulus patterns. The trained UDE accurately predicts responses to pulse stimuli (not seen during training with step stimulus). Vertical line at t=60s marks end of training data, showing successful extrapolation. This predictive capability enables assessment of real-world environments before exposure.

Priority directions include multi-site clinical validation with 100+ autistic participants comparing UDE predictions to behavioral assessments; implementation of hierarchical UDEs modulating parameters based on cognitive state; development of multi-modal sensory models with cross-modal interactions; integration with predictive coding frameworks connecting our model to computational theories of autism [16, 27, 13]; population-level analysis revealing sensory subtypes through parameter clustering; and closed-loop intervention systems automatically adjusting environmental parameters based on predicted risk.

Universal Differential Equations successfully bridge mechanistic understanding and data-driven flexibility to model sensory processing in autism. By augmenting neurophysiological models with neural networks only where knowledge is incomplete, UDEs achieve 90.8% improvement over pure machine learning while maintaining interpretability, recovering physiological parameters within 2% error, and providing quantitative risk assessment (17.2% risk for pulse stimuli).

This work establishes foundations for personalized, evidence-based interventions previously impossible with traditional approaches. The ability to predict individual responses enables proactive support strategies from environmental modifications to wearable alert systems. As we advance toward precision medicine in neurodevelopmental conditions, frameworks that balance interpretability with personalization become increasingly critical, demonstrating that we need not choose between understanding mechanisms and predicting outcomes.

Future deployment could transform sensory support in autism from reactive crisis management to proactive prevention. By providing personalized quantitative assessments, this framework empowers autistic individuals, informs caregivers, and guides evidence-based accommodations. The success in this challenging domain suggests broader applicability to other neurological conditions where individual variability challenges traditional modeling.

This technology requires careful ethical consideration to empower rather than impose neurotypical standards, using assessments to inform accommodation rather than exclusion. Individual parameters

must be protected as sensitive health information while ensuring accessibility regardless of socioeconomic resources. The framework enables self-advocacy through quantitative evidence of sensory needs, reducing emotional labor while improving understanding among educators, employers, and families through objective validation of sensory differences.

## References


[1] Lord, C., Brugha, T. S., Charman, T., Cusack, J., Dumas, G., Frazier, T., ... & Veenstra-VanderWeele, J. (2020). Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1), 1-23.

[2] American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Arlington, VA: American Psychiatric Publishing.

[3] Maenner, M. J., Shaw, K. A., Baio, J., et al. (2020). Prevalence of autism spectrum disorder among children aged 8 years. *MMWR Surveillance Summaries*, 69(4), 1-12.

[4] Robertson, C. E., & Baron-Cohen, S. (2017). Sensory perception in autism. *Nature Reviews Neuroscience*, 18(11), 671-684.

[5] Tomchek, S. D., & Dunn, W. (2007). Sensory processing in children with and without autism. *American Journal of Occupational Therapy*, 61(2), 190-200.

[6] Leekam, S. R., Nieto, C., Libby, S. J., Wing, L., & Gould, J. (2007). Describing the sensory abnormalities of children and adults with autism. *Journal of Autism and Developmental Disorders*, 37(5), 894-910.

[7] Ben-Sasson, A., Hen, L., Fluss, R., Cermak, S. A., Engel-Yeger, B., & Gal, E. (2009). A meta-analysis of sensory modulation symptoms in individuals with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 39(1), 1-11.

[8] Ashburner, J., Ziviani, J., & Rodger, S. (2008). Sensory processing and classroom emotional, behavioral, and educational outcomes in children with autism spectrum disorder. *American Journal of Occupational Therapy*, 62(5), 564-573.

[9] Marco, E. J., Hinkley, L. B., Hill, S. S., & Nagarajan, S. S. (2011). Sensory processing in autism: a review of neurophysiologic findings. *Pediatric Research*, 69(8), 48-54.

[10] Hazen, E. P., Stornelli, J. L., O'Rourke, J. A., Koesterer, K., & McDougle, C. J. (2014). Sensory symptoms in autism spectrum disorders. *Harvard Review of Psychiatry*, 22(2), 112-124.

[11] Green, S. A., Hernandez, L., Bookheimer, S. Y., & Dapretto, M. (2015). Reduced modulation of thalamocortical connectivity during exposure to sensory stimuli in ASD. *Autism Research*, 8(1), 94-103.

[12] Brandwein, A. B., Foxe, J. J., Butler, J. S., et al. (2013). The development of multisensory integration in high-functioning autism. *Frontiers in Human Neuroscience*, 7, 267.

[13] Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16(10), 504-510.

[14] Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, 2(5), 255-267.

[15] Gogolla, N., LeBlanc, J. J., Quast, K. B., Südhof, T. C., Fagiolini, M., & Hensch, T. K. (2009). Common circuit defect of excitatory-inhibitory balance in mouse models of autism. *Journal of Neurodevelopmental Disorders*, 1(2), 172-181.

[16] Lawson, R. P., Rees, G., & Friston, K. J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, 8, 302.

[17] Verhoog, M. B., Goriounova, N. A., Obermayer, J., et al. (2013). Mechanisms underlying the rules for associative plasticity at adult human neocortical synapses. *Journal of Neuroscience*, 33(43), 17197-17208.

[18] Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1), 1-24.

[19] Chen, R. T., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In *Advances in Neural Information Processing Systems* (pp. 6571-6583).

[20] Dupont, E., Doucet, A., & Teh, Y. W. (2019). Augmented neural ODEs. In *Advances in Neural Information Processing Systems* (pp. 3140-3150).

[21] Kidger, P., Morrill, J., Foster, J., & Lyons, T. (2020). Neural controlled differential equations for irregular time series. In *Advances in Neural Information Processing Systems*, 33, 6696-6707.

[22] Massaroli, S., Poli, M., Park, J., Yamashita, A., & Asama, H. (2020). Dissecting neural ODEs. In *Advances in Neural Information Processing Systems*, 33, 3952-3963.

[23] Rackauckas, C., Ma, Y., Martensen, J., Warner, C., Zubov, K., Supekar, R., ... & Edelman, A. (2020). Universal differential equations for scientific machine learning. *arXiv preprint arXiv:2001.04385*.

[24] Ma, Y., Dixit, V., Innes, M., Guo, X., & Rackauckas, C. (2021). A comparison of automatic differentiation and continuous sensitivity analysis for derivatives of differential equation solutions. *arXiv preprint arXiv:2012.06684*.

[25] Yazdani, A., Lu, L., Raissi, M., & Karniadakis, G. E. (2020). Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS Computational Biology*, 16(11), e1007575.

[26] Santana, S., Martius, G., & Moreno-Muñoz, P. (2021). Universality and individuality in neural dynamics across large populations. *Nature Neuroscience*, 24(1), 123-134.

[27] Van de Cruys, S., Evers, K., Van der Hallen, R., Van Eylen, L., Boets, B., de-Wit, L., & Wagemans, J. (2014). Precise minds in uncertain worlds: Predictive coding in autism. *Psychological Review*, 121(4), 649.

[28] Rackauckas, C., Innes, M., Ma, Y., Bettencourt, J., White, L., & Dixit, V. (2019). DiffEqFlux.jl - A Julia library for neural differential equations. *arXiv preprint arXiv:1902.02376*.

[29] Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1), 65-98.

[30] Katz, B., & Miledi, R. (1967). The role of calcium in neuromuscular facilitation. *The Journal of Physiology*, 195(2), 481-492.

[31] Torre, V., Ashmore, J. F., Lamb, T. D., & Menini, A. (1995). Transduction and adaptation in sensory receptor cells. *Journal of Neuroscience*, 15(12), 7757-7768.

[32] Zucker, R. S., & Regehr, W. G. (2002). Short-term synaptic plasticity. *Annual Review of Physiology*, 64(1), 355-405.

[33] Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, 275(5297), 221-224.

[34] Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends in Cognitive Sciences*, 13(7), 293-301.

[35] Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, 2(1), 79-87.

[36] Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neurobiology*, 17(4), 423-429.

[37] De Palo, G., Facchetti, G., Mazzolini, M., Menini, A., Torre, V., & Altafini, C. (2012). Common dynamical features of sensory adaptation in photoreceptors and olfactory sensory neurons. *Scientific Reports*, 3, 1251.

[38] Roberts, T. P., Khan, S. Y., Rey, M., Monroe, J. F., Cannon, K., Blaskey, L., ... & Edgar, J. C. (2010). MEG detection of delayed auditory evoked responses in autism spectrum disorders. *NeuroReport*, 21(4), 259-263.

[39] Edgar, J. C., Khan, S. Y., Blaskey, L., et al. (2015). Neuromagnetic oscillations predict evoked-response latency delays and core language deficits in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 45(2), 395-405.

[40] Sohal, V. S., Zhang, F., Yizhar, O., & Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature*, 459(7247), 698-702.

[41] Yizhar, O., Fenno, L. E., Prigge, M., et al. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, 477(7363), 171-178.

[42] Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, 13(1), 51-62.

[43] Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience*, 14(6), 775-782.

[44] Uljarević, M., & Hamilton, A. (2017). Recognition of emotions in autism: A formal meta-analysis. *Journal of Autism and Developmental Disorders*, 43(7), 1517-1526.

[45] Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience Bulletin*, 33(2), 183-193.

[46] Silver, R. A. (2010). Neuronal arithmetic. *Nature Reviews Neuroscience*, 11(7), 474-489.

[47] Louizos, C., Welling, M., & Kingma, D. P. (2017). Learning sparse neural networks through L0 regularization. *arXiv preprint arXiv:1712.01312*.

[48] Gross, J., Baillet, S., Barnes, G. R., et al. (2013). Good practice for conducting and reporting MEG research. *NeuroImage*, 65, 349-363.

[49] Baillet, S. (2017). Magnetoencephalography for brain electrophysiology and imaging. *Nature Neuroscience*, 20(3), 327-339.

[50] Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). *Neuronal dynamics: From single neurons to networks and models of cognition*. Cambridge University Press.

[51] Dayan, P., & Abbott, L. F. (2005). *Theoretical neuroscience: Computational and mathematical modeling of neural systems*. MIT Press.

[52] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[53] Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of Adam and beyond. *arXiv preprint arXiv:1904.09237*.

[54] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., & Mishchenko, E. F. (1962). *The mathematical theory of optimal processes*. Interscience Publishers.

[55] Tavassoli, T., Miller, L. J., Schoen, S. A., Nielsen, D. M., & Baron-Cohen, S. (2014). Sensory over-responsivity in adults with autism spectrum conditions. *Autism*, 18(4), 428-432.

[56] Schoen, S. A., Miller, L. J., Brett-Green, B. A., & Nielsen, D. M. (2009). Physiological and behavioral differences in sensory processing: A comparison of children with autism spectrum disorder and sensory modulation disorder. *Frontiers in Integrative Neuroscience*, 3, 29.

[57] Greydanus, S., Dzamba, M., & Yosinski, J. (2019). Hamiltonian neural networks. In *Advances in Neural Information Processing Systems*, 32.

[58] Cranmer, M., Greydanus, S., Hoyer, S., et al. (2020). Lagrangian neural networks. *arXiv preprint arXiv:2003.04630*.

[59] Delorme, A., Platt, G., Tager-Flusberg, H., & Nelson, C. A. (2013). Enhanced detection of artifacts in EEG data using independent component analysis and extended infomax. *NeuroImage*, 34(4), 1443-1449.

[60] Keehn, B., Müller, R. A., & Townsend, J. (2013). Atypical attentional networks and the emergence of autism. *Neuroscience & Biobehavioral Reviews*, 37(2), 164-183.

[61] Baum, S. H., Stevenson, R. A., & Wallace, M. T. (2015). Behavioral, perceptual, and neural alterations in sensory and multisensory function in autism spectrum disorder. *Progress in Neurobiology*, 134, 140-160.

[62] Brandwein, A. B., Foxe, J. J., Butler, J. S., et al. (2015). Neurophysiological indices of atypical auditory processing and multisensory integration are associated with symptom severity in autism. *Journal of Autism and Developmental Disorders*, 45(1), 230-244.

[63] Vazquez, M., et al. (2024). Comparative analysis of hyperparameter selection in scientific machine learning. *Journal of Computational Physics*, 475, 111845.