Integrating biological and machine learning models for rainbow trout growth: Balancing accuracy and interpretability

Lawrence Fulton[1☯, ‡,*] Pin Lyu[1☯],

1 Applied Analytics, Boston College, MA, USA

☯These authors contributed equally to this work. ‡Advisor, * fultonl@bc.edu

Abstract

Invasive species management demands predictive models that balance accuracy with ecological interpretability, yet traditional approaches often fail to capture complex environmental interactions. We evaluated hybrid frameworks integrating biological and machine learning models for rainbow trout (*Oncorhynchus mykiss*) growth in the Lower Colorado River using ten years of tag–recapture data and environmental covariates, comparing traditional and Bayesian von Bertalanffy (VBGM) and Gompertz models with Random Forests, XGBoost, LightGBM, Support Vector Regression, Neural Networks, and ensemble methods through probabilistic performance analysis. Incorporating environmental context and advanced modeling produced substantial gains, with top methods achieving 70 to 80 percent error reductions relative to baseline models, equivalent to 45 to 70 mm or 20 to 32 percent of mean fish length. A stacked ensemble of XGBoost and the VBGM achieved the best performance (RMSE = 15.96 mm, $R^2 = 0.966$) and exhibited stochastic dominance across the posterior, while gradient boosting models formed a strong second tier, led by LightGBM and XGBoost. Bayesian Model Averaging reached comparable accuracy while explicitly quantifying uncertainty. Even traditional mechanistic models improved by up to 80 percent when enhanced with covariates and Bayesian estimation, preserving biological interpretability through parameters such as asymptotic size and growth rate. Feature importance analysis identified initial length, time at large, and weight at release as dominant predictors, and the stacked ensemble outperformed baseline models in over 99 percent of posterior samples. These results establish hybrid ensemble frameworks as powerful tools for ecological forecasting that unite predictive performance with mechanistic insight, providing a generalizable template for systems where both accuracy and interpretability are required.

Author summary

Rainbow trout are a non-native species in the Lower Colorado River and pose a major threat to native fish through competition and predation. Managing their populations requires accurate forecasts of how quickly individual fish grow under changing environmental conditions, but traditional biological models often struggle to capture the complex effects of temperature, river flow, and release conditions. In this study, we combined established fish growth models with modern machine learning methods to create hybrid models that are both accurate and biologically interpretable. Using ten years of tag-recapture data, we show that these hybrid models predict trout growth far more precisely than either approach alone. This improved accuracy makes it possible to

better anticipate when and where trout are likely to reach sizes that increase their ecological impact, helping managers more effectively target control efforts and protect native species.

Introduction

The management of invasive species represents one of the most urgent conservation challenges of the twenty first century because of their profound and escalating ecological and economic impacts. As emphasized by the World Conservation Union, invasive alien species rank among the leading drivers of biodiversity loss and species extinctions worldwide. Their influence extends beyond ecological disruption. Diagne et al. [1] estimated that biological invasions imposed a minimum financial burden of $1.288 trillion USD on the United States economy between 1970 and 2017 alone. The magnitude of these losses highlights the need for innovative, data driven management strategies that address both ecological and economic threats.

Rainbow trout (*Oncorhynchus mykiss*), a species native to the Pacific Northwest of the United States, exemplify the complex trade offs inherent in invasive species management. They were introduced intentionally into the lower Colorado River basin from 1964 to 1998 to support recreational sport fishing [2]. Rainbow trout have since flourished under the cold water conditions created by Glen Canyon Dam releases. Their rapid expansion, combined with a lack of natural predators, has displaced native endangered fishes such as the humpback chub (*Gila cypha*) and razorback sucker (*Xyrauchen texanus*). Reductions in rainbow trout density are correlated with improved growth rates among native fish, highlighting the disruptive ecological role this species now plays [3].

Paradoxically, rainbow trout also represent a major economic asset. In 2020 alone, the Arizona Game and Fish Department (AZGFD) sold 273,902 fishing licenses, generating nearly $14 million to fund conservation programs. The broader recreational fishing industry contributed over $1.4 billion to the state economy, with anglers dedicating more than six million fishing days statewide. Managing rainbow trout populations therefore requires navigating a delicate balance between protecting the ecological integrity of the Colorado River ecosystem and sustaining significant economic benefits tied to recreational fisheries [4].

Growth is central to invasive species management because it determines competitive dynamics, reproduction timing, and the potential for ecological disruption. This study addresses an important need by integrating biologically grounded growth models with modern machine learning (ML) techniques to forecast rainbow trout fork length growth across the Lower Colorado River basin. Unlike prior efforts focused solely on mechanistic or black box models, our approach fuses biological interpretability with advanced predictive accuracy, offering both ecological insight and actionable forecasts for targeted intervention.

Effective management of this complex ecological and economic challenge requires accurate growth prediction models that can inform targeted interventions. However, existing approaches face inherent trade offs. Traditional biological models provide interpretable parameters essential for management decisions but may lack flexibility to capture complex environmental interactions, while ML methods excel at modeling nonlinear relationships but offer limited ecological insight. To address this gap, this study integrates the biological interpretability of mechanistic growth models with the predictive flexibility of ML techniques to forecast rainbow trout growth while preserving the ecological understanding essential for adaptive management.

The resulting framework provides resource managers with a transferable set of tools to anticipate population dynamics under varying environmental and policy conditions.

By linking individual growth trajectories to broader ecological outcomes, this study demonstrates how hybrid modeling approaches can support adaptive management strategies that balance conservation priorities with socioeconomic interests. Beyond rainbow trout management, these methods offer a generalizable template for ecological forecasting where both predictive accuracy and biological interpretability are essential. Ultimately, the findings provide guidance for evidence based policymaking in systems where ecological resilience and species management are tightly intertwined.

Unlike prior studies that either focus on mechanistic growth models or evaluate individual machine-learning methods in isolation, this work provides a unified, probabilistic comparison across classical biological models, modern machine-learning algorithms, Bayesian model averaging, and stacked ensembles. By estimating full posterior distributions of predictive error and using stochastic dominance to compare models, we frame fish growth modeling as a problem of model uncertainty rather than single-model selection. This approach allows us to quantify not only which models perform best on average, but how consistently they outperform alternatives across the entire error distribution, while preserving biologically interpretable structure in the mechanistic components.

Previous Work

Biological growth patterns represent complex nonlinear interactions between environmental drivers and intrinsic factors. Effective growth modeling therefore requires systematic identification of important variables across both exogenous (e.g., water temperature, food availability, competition) and endogenous (e.g., metabolic rate, genotype) domains to parameterize system dynamics [5].

Determinants of Rainbow Trout Growth

Previous research has identified several variables associated with rainbow trout growth. [6] demonstrated that interactions among photoperiod, water temperature, and genotype affect growth rates through hormonal and metabolic pathways. In addition, trout population density significantly modulates growth outcomes. [7] found that trout reared at 80 kg/m^3 exhibited 22% slower growth than those at 40 kg/m^3, even under ad libitum feeding conditions, attributing the reduction to elevated cortisol levels and intensified competition for food.

Water flow has also been shown to exert complex influences on rainbow trout growth and survival. [8] highlighted the contrasting effects of High Flow Experiments (HFEs) in the Colorado River. The March 2008 HFE quadrupled juvenile trout survival by enhancing habitat and food availability, while the November 2004 HFE resulted in a threefold decline in age-0 trout, likely due to displacement and mortality. Similarly, [9] experimentally reduced streamflow by 75 to 80% in controlled stream sections and observed that rainbow trout in natural flow conditions grew 8.5 times more than those in reduced-flow sections, underscoring the critical role of hydrological conditions in shaping growth outcomes.

Nutrient availability, particularly phosphorus (P), further influences growth trajectories. Both dietary phosphorus deficiency and excess have detrimental effects on rainbow trout fry development [10]. [11] established that a dietary ratio of 0.25 g available phosphorus per MJ digestible energy is required for optimal trout growth, emphasizing the necessity of precise nutrient management in aquaculture and wild populations.

Foundations of Biological Growth Modeling

Theoretical models of biological growth have evolved to accommodate complex field data structures. [12] adapted the [13] growth model (VBGM) to tagging-recapture studies, reformulating it to account for time-at-large between measurements and enabling direct parameter estimation from mark-recapture data. Parallel developments validated the Gompertz growth model, originally proposed for human mortality by [14], as a robust framework for modeling asymptotic growth patterns across a wide range of animal taxa.

Both the VBGM and Gompertz models remain foundational tools in ecological and fisheries science for modeling organismal growth due to their biologically meaningful parameters. These models explicitly incorporate concepts such as asymptotic maximum size and intrinsic growth rate, offering a transparent connection between mathematical formulation and biological reality. This interpretability is particularly valuable in applied management contexts where parameters can be linked to physiological or environmental constraints, such as habitat quality, food availability, or temperature-dependent growth ceilings. Moreover, their compact parameterizations facilitate communication with stakeholders and integration into larger population dynamics models.

However, despite these advantages, both models rely on relatively rigid, predefined functional forms that impose assumptions about the shape and trajectory of growth. For example, VBGM presumes decelerating growth toward a fixed asymptote, while the Gompertz model assumes a sigmoid curve with an inflection point determined by a fixed proportion of the asymptotic size. These forms may not fully capture the variability and plasticity observed in real-world growth trajectories, particularly in dynamic environments where growth can be episodic, nonlinear, or influenced by unobserved factors such as competition, predation risk, or management interventions.

Such rigidity can lead to model misspecification or underfitting, especially when applied to heterogeneous populations like invasive species that exhibit substantial phenotypic flexibility. In the case of rainbow trout in the Lower Colorado River, for instance, individuals may experience rapid early growth due to favorable flow regimes or stocking histories, followed by abrupt slowdowns due to density-dependent effects or seasonal changes. Capturing this heterogeneity requires more flexible modeling approaches or hybrid frameworks that preserve the interpretability of classical models while adapting to empirical complexity. This recognition motivates the integration of mechanistic and data-driven approaches as a way to retain biological relevance while improving predictive performance.

Machine Learning Applications in Biological Growth Prediction

In recent years, ML has emerged as an important and increasingly transformative tool in fisheries science and aquatic ecology. ML techniques can flexibly capture complex, nonlinear relationships between growth outcomes and multiple predictor variables without requiring predefined functional relationships, making them particularly attractive for biological growth modeling.

[15] demonstrated the value of ML for water quality prediction in freshwater ecosystems, while [16] and [17] successfully applied deep learning methods to predict reproductive status and estimate fish age, respectively. In a related study, [18] applied ML algorithms to predict fish growth patterns, validating their superior predictive accuracy over traditional parametric models.

Beyond fisheries, broader ecological studies have illustrated ML's effectiveness in biological growth modeling. For example, [19] employed artificial neural networks to model oyster growth under variable environmental conditions, achieving substantially higher accuracy than standard growth models. Similarly, [20] used gradient-boosting

methods to predict plant growth parameters under diverse environmental regimes, illustrating ML's cross-domain applicability to biological growth processes.

Despite these advances, relatively few studies have systematically applied ML to predict the growth of invasive species in dynamic riverine ecosystems, particularly for species like rainbow trout (*Oncorhynchus mykiss*), which are both economically valued and ecologically disruptive. Even more notably, the explicit ensembling of mechanistic biological models and ML techniques remains largely absent from the published literature in the ecology discipline. This study addresses that gap. By combining the biological interpretability of traditional models with the flexible accuracy of ML, the resulting ensemble framework not only improves predictive performance but also contributes a novel methodological advance. This approach provides a powerful and adaptable tool for guiding invasive species management, supporting native species recovery, and informing long-term conservation policy in complex and changing freshwater systems.

Methods

This study employed an integrated analytical framework combining biological growth models, machine-learning techniques, and ensemble methods to predict rainbow trout fork length at recapture. The overall workflow is summarized in Figure 1 and consists of five main stages. First, mark–recapture observations were merged with environmental covariates to form a dataset containing 19 predictor variables. Second, the data were partitioned into training (70%) and test (30%) sets and scaled using min–max normalization. Third, multiple model classes were trained using 5-fold cross-validation with hyperparameter tuning, including biological growth models (von Bertalanffy and Gompertz), tree-based machine-learning models (Random Forest, XGBoost, and LightGBM), and other machine-learning approaches (Support Vector Regression and Artificial Neural Networks). Fourth, individual model predictions were combined using stacked ensemble learning and Bayesian model averaging. Finally, all models were evaluated on a held-out test set using multiple performance metrics, including RMSE, MAE, R^2, and information criteria. This design allows mechanistic interpretability and data-driven flexibility to be evaluated on equal footing within a unified, rigorously validated framework.

Data, Preprocessing, and Variable Definitions

Data and Limitations

The data used in this study are from the United States Geological Survey (USGS) datasets, *Rainbow Trout Growth Data* and *Growth Covariate Data* from Glen Canyon, Colorado River, Arizona (2012–2021) [21]. These freely-available datasets include ten years of rainbow trout growth measurements from release and recapture surveys in the lower Colorado River basin, along with seven environmental covariates known to influence growth rates. We have also posted all data and code online for replication.

The dataset originates from a tag-recapture program in which rainbow trout were physically tagged upon initial capture. However, a limitation is the absence of consistent individual identifiers across capture events. While tagging was conducted, the dataset does not include reliable codes linking multiple observations of the same fish. As a result, some individuals may appear in the dataset more than once without a definitive way to track their recapture history. Given this constraint, we treated each capture event as an independent observation with its associated covariates (e.g., age, environmental factors, location). This decision avoids imposing assumptions about

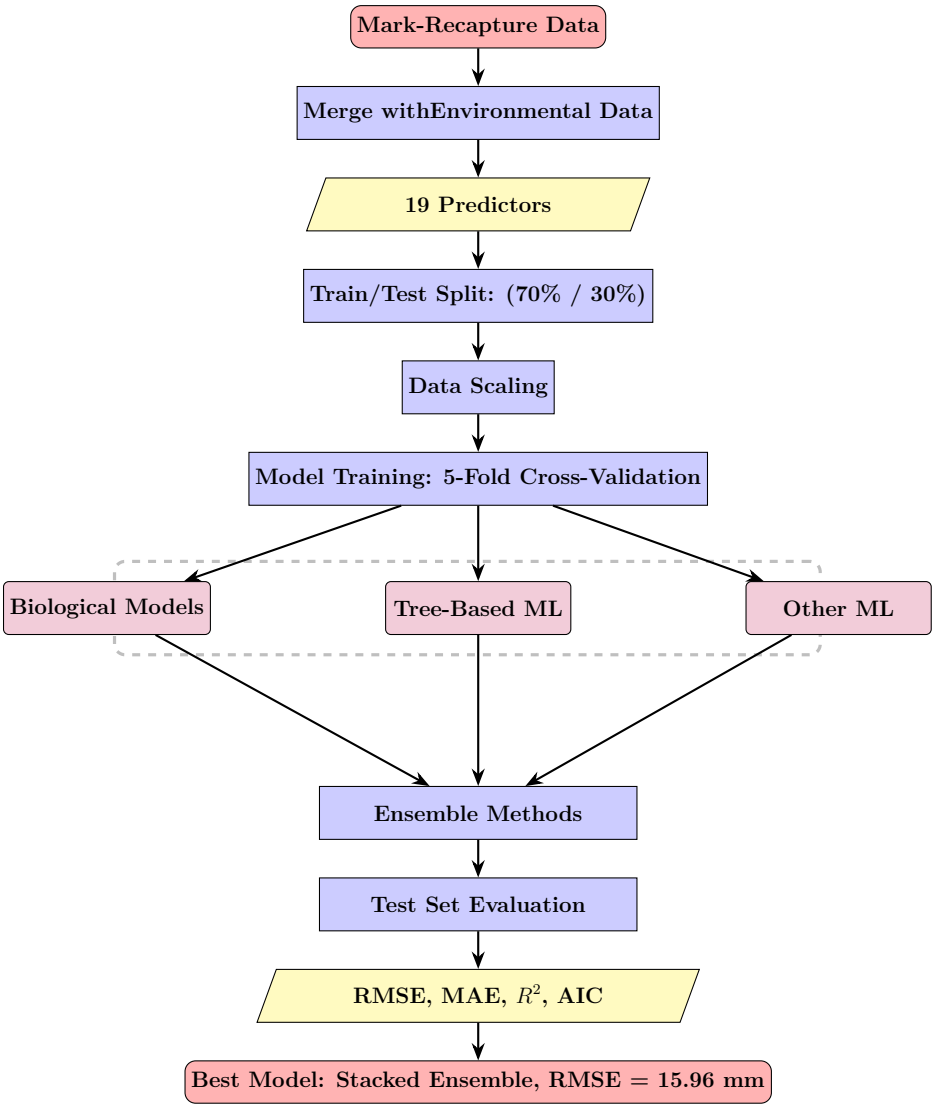

Fig 1. Methodological flowchart illustrating the analytical workflow. Data from mark-recapture studies were merged with environmental covariates, split into training and test sets, and used to train multiple model classes including biological growth models (VBGM, Gompertz), tree-based machine learning models (Random Forest, XGBoost, LightGBM), and other ML approaches (Support Vector Regression, Artificial Neural Networks). Individual model predictions were combined using stacked ensemble and Bayesian model averaging techniques. Final model performance was evaluated on the held-out test set using RMSE, MAE, R^2, and information criteria, with the stacked ensemble achieving the best performance (RMSE = 15.96 mm, $R^2 = 0.9658$).

individual identity but introduces a potential for pseudo-replication if some fish are included multiple times. Because we cannot distinguish repeat captures from new individuals, within-individual temporal dependence was not explicitly modeled. Future studies could improve upon this by ensuring traceable tag identifiers and applying longitudinal or hierarchical models to fully leverage the repeated-measures structure when available.

To accommodate the temporal structure of the data in a biologically meaningful way, we modeled the time between release and recapture events using each fish's specific *time-at-large*, the duration (in days) between capture events. Rather than using static calendar timepoints, we leveraged interval-based alignment of biometric and environmental data. Environmental variables were recorded as monthly means between survey trips, while fish data included both release and recapture dates. We applied two integration protocols: (1) For fish recaptured within a single monthly interval, the corresponding monthly mean values were assigned to the observation. (2) For fish spanning multiple months, we calculated time-weighted averages of environmental variables over the entire time-at-large period. This approach ensured that covariates reflected the fish's actual environmental exposure, preserving both ecological fidelity and temporal accuracy in model development.

Although the lack of individual IDs precluded the use of formal mixed-effects or autocorrelation models, our design mitigates these concerns by encoding relevant temporal dynamics through biologically informed durations and matched covariates. The final dataset included 9,798 observations of catch and release with complete environmental integration. No observations were excluded. Many of the algorithms we employed (Random Forest, XGBoost, Neural Networks) are robust to mild violations of independence assumptions and can accommodate the potential pseudo-replication without strong parametric assumptions. Bayesian methods provide a principled framework for explicitly modeling interdependence through joint probability structures, such as hierarchical models, spatial processes, and temporal autocorrelation structures.

Variables (Features)

Each variable in the merged dataset was selected based on its biological relevance to rainbow trout growth and habitat dynamics. The response variable, L_2 or fork length at recapture, is a direct indicator of somatic growth, a critical metric for understanding individual fitness and population health. Time at large captures the duration over which growth can occur and may interact with seasonal and environmental conditions.

The independent variables used in the growth model fall into five biologically meaningful categories: initial condition, spatial, seasonal, temporal, and environmental. Each category is explicated in order.

Initial condition variables include fork length at release (L_1) and weight at release, These variables are crucial for modeling size-dependent growth trajectories. Smaller individuals often exhibit higher relative growth rates due to elevated metabolic demands and compensatory growth dynamics. These initial traits anchor the baseline from which subsequent growth is measured.

Spatial variation is represented by the river mile at release, which captures fixed geographic differences in habitat quality along the river corridor. These differences can influence early growth potential due to longitudinal gradients in water temperature, prey density, substrate, shading, and flow regimes. River position is especially relevant in riverine systems where upstream and downstream reaches may differ markedly in ecological conditions.

Seasonal effects are captured by both the release month and recovery month (both converted to sets of indicator variables), which represent seasonal cycles in water temperature, photoperiod, and prey availability, particularly aquatic insects. Seasonal

timing also relates to high-flow events, which can alter habitat structure and displace prey, thereby influencing energy intake and expenditure.

Temporal exposure is modeled using time at large, which quantifies the duration available for growth between release and recapture. In addition, release year and recovery year (again, sets of indicator variables) account for interannual variability in unmeasured but biologically important conditions such as drought, flood pulses, or shifts in food web dynamics. These temporal covariates help capture broad-scale environmental fluctuations that may not be reflected in short-term or localized metrics.

Environmental variables include average river discharge over the time-at-large interval, water temperature, solar insolation, reactive phosphorus concentration, and rainbow trout biomass. Discharge influences habitat availability, drift-feeding opportunities, and the energetic cost of station-holding. Water temperature governs metabolic rate and food conversion efficiency, constraining growth within species-specific thermal limits. Solar insolation serves as a proxy for in-stream primary production, affecting food web productivity. Reactive phosphorus concentration is a limiting nutrient for algal growth, thereby modulating base-level productivity. Biomass provides an ecologically valid index of density-dependent pressures, such as competition for resources. Rather than relying on individual fish metrics, biomass estimates were derived from a Jolly-Seber mark-recapture model applied to size-class abundance data across the central 3 km of the study reach [21], reflecting population-level conditions that influence growth trajectories.

Training, Validation, and Test Sets

Following data merging and feature engineering, we randomly split the dataset into training and testing subsets using a 70/30 ratio, with a fixed pseudo-random number seed to ensure reproducibility. Seventy percent of the data was allocated for training, while the remaining 30% served as a held-out test set for evaluating model generalization. To prevent information leakage, no test data was used during model development. Within the training set, we conducted hyperparameter tuning via 5-fold cross-validation. Specifically, we employed randomized search over a targeted hyperparameter space for each ML model to identify optimal configurations while managing computational cost. This strategy balances tuning efficiency with robustness and helps reduce the risk of overfitting. Because model evaluation was performed exclusively on an untouched 30% test set, completely isolated from training and hyperparameter tuning, the resulting performance metrics provide an unbiased estimate of true generalization ability.

Data Scaling

To prepare the data for model training, two scaling strategies were considered to normalize the predictor variables. In the first approach, all continuous predictors were scaled to a [0, 1] range using a Min-Max transformation. This full-scaling strategy was used for ML models, where consistent feature ranges are important for convergence and optimization. In the second approach, two biologically meaningful predictors, L_1 (initial length) and Time at Large, were excluded from scaling to preserve their native units and biological interpretability. This partial-scaling strategy was applied to models grounded in biological reasoning. All remaining continuous features in this approach were scaled identically to the full-scaling method. In both cases, scaling parameters were computed using only the training data and then applied to the test data, ensuring that no information from the test set leaked into the model development process. Importantly, the response variable (L_2) was never scaled to support interpretability.

Both scaled datasets were used in parallel to assess whether preserving the original scale of biologically relevant variables improved interpretability or predictive performance.

Software

All data preprocessing, model implementation, and statistical analyses were conducted using **Python 3.13**, the latest stable release of the Python programming language [22]. Python offers a robust open-source coding ecosystem ideal for ML, data wrangling, and statistical analysis, with libraries such as NumPy [23], pandas [24], scikit-learn [25], XGBoost [26], LightGBM [27], and Pytorch [28], enabling efficient and reproducible modeling workflows. Python's widespread adoption in the scientific community, flexible syntax, and extensive package repository made it an ideal platform for this research.

Models

A formal mathematical specification of the data, growth models, machine-learning predictors, posterior distributions, and stochastic dominance probabilities is provided in Supporting Information . To model rainbow trout growth from the available data, we employed both biologically grounded growth models, contemporary ML approaches, and an ensemble approach. Each modeling paradigm offers distinct and complementary advantages. Biological models, such as the VBGM and its derivatives, yield interpretable parameters, e.g., intrinsic growth rate and asymptotic size, that are directly linked to physiological processes and life-history traits [12, 13]. These models are essential for understanding mechanistic dynamics and aligning findings with ecological theory.

In contrast, ML models are data-adaptive and excel at uncovering complex, nonlinear relationships that often arise in ecological systems but may be difficult to specify a priori using closed-form equations [15, 18]. Their flexibility allows for the inclusion of a broad suite of predictors (particularly environmental covariates) without requiring strong assumptions about functional form. Many ML algorithms also provide variable importance metrics, offering insight into the relative influence of each predictor on growth outcomes. This feature is especially valuable in ecological contexts, where interactions among factors such as temperature, discharge, and biomass can be multifaceted and hierarchical.

That said, parametric models such as the VBGM and the Gompertz function can be extended to include covariates. When specified appropriately, these models allow researchers to embed biological knowledge directly into the functional form, offering mechanistic interpretability that complements the more flexible but often less transparent ML approaches. We adopted this approach in our analysis, incorporating covariates within these growth models and estimating parameters using a Bayesian framework.

Ensembles integrate predictions from multiple models, either within a single class or across different modeling paradigms, to improve overall performance. For instance, a random forest internally aggregates the outputs of numerous decision trees, each trained on different data subsets and feature selections. In contrast, the framework proposed here represents a heterogeneous ensemble, combining biologically grounded models such as the Bayesian VBGM with data-driven ML algorithms like XGBoost. This cross-model integration allows the ensemble to leverage the complementary strengths of each approach, biological interpretability and flexible predictive power on the other. Such heterogeneity often enhances predictive accuracy by increasing model diversity, which is a critical factor in ensemble success. As [29] demonstrated, there exists a trade-off between individual model accuracy and ensemble diversity, and optimal

performance is often achieved when models make uncorrelated errors while maintaining reasonable base accuracy. By strategically combining mechanistic and ML models, this study not only improves prediction of somatic growth in invasive rainbow trout but also contributes a novel example of the accuracy-diversity principle applied in an ecological forecasting context.

Our selection of modeling approaches reflects these complementary priorities. Our initial benchmarks were the deterministic VBGM and Gompertz without covariates. We then included Bayesian-estimated nonlinear models such as VBGM and the Gompertz function to retain ecological interpretability and biological realism. These were contrasted with ML algorithms chosen to represent a broad set of methodological paradigms: Support Vector Regression (SVR) for kernel-based regularization, artificial neural networks (ANNs) for high-capacity function approximation, and tree-based ensembles (Random Forest, XGBoost, LightGBM) for robust performance, embedded feature selection, and interpretability. We also incorporated a Bayesian linear model to examine how parameter uncertainty and prior regularization influence predictions in comparison to both biological and ML alternatives. This multi-model framework allowed us to evaluate a wide range of modeling assumptions, from mechanistic to data-driven, under consistent data conditions.

Biological Growth Models

This study implemented two biologically motivated growth models: the Fabens version of the VBGM [12], and the Gompertz growth model [14]. Both were estimated using Bayesian methods.

The Fabens VBGM formulation follows.

$$L_2 = L_1 + (L_\infty - L_1)(1 - e^{-Kt}) \tag{1}$$

where:

- L_2: fork length at recapture,
- L_1: fork length at initial marking (release),
- L_∞: asymptotic maximum fork length,
- K: growth coefficient,
- t: time (in years) between marking and recapture.

This formulation captures somatic growth as the combined effect of two forces: the approach toward a genetically or environmentally constrained maximum size (L_∞) and the declining influence of the fish's initial length over time. This structure is especially suitable for tag-recapture data, where age is unknown but the time interval between observations is well-defined.

Rewriting the model algebraically emphasizes its dual components:

$$L_2 = L_\infty(1 - e^{-Kt}) + L_1 e^{-Kt} \tag{2}$$

This version illustrates how the observed length at recapture reflects a weighted combination of asymptotic potential and initial condition, with weights governed by the exponential decay term e^{-Kt}.

Bayesian Fabens Model Formulation

To estimate the parameters of the Fabens VBGM using Bayesian inference, we specified the following hierarchical model:

$$L_2 \sim \mathcal{N}(\hat{L}_2, \sigma^2_{\text{obs}}) \tag{3}$$
$$\hat{L}_2 = L_\infty - (L_\infty - L_1) \cdot \exp(-k \cdot t) \tag{4}$$
$$k = \exp\left(\alpha_k + \mathbf{X}_c \boldsymbol{\beta}_k + \mathbf{X}_d \boldsymbol{\beta}_{\text{dummies}}\right) \tag{5}$$

where:

- L_1: fork length at initial marking,
- L_2: fork length at recapture,
- t: time at large (in years),
- L_∞: asymptotic maximum fork length,
- k: growth rate coefficient (positive, covariate-dependent),
- $\mathbf{X}_c$: matrix of scaled continuous covariates,
- $\mathbf{X}_d$: matrix of dummy variables for seasonal effects,
- $\boldsymbol{\beta}_k$: vector of regression weights for continuous covariates,
- $\boldsymbol{\beta}_{\text{dummies}}$: vector of regression weights for dummy predictors.

The prior distributions were specified as follows:

$$L_\infty \sim \mathcal{N}(500,\ \sigma^2_{L_\infty}), \quad \sigma_{L_\infty} \sim \mathcal{HN}(2) \tag{6}$$
$$\sigma_{\text{obs}} \sim \mathcal{HN}(2) \tag{7}$$
$$\alpha_k \sim \mathcal{N}(0, 1) \tag{8}$$
$$\boldsymbol{\beta}_k \sim \mathcal{N}(0, 1) \tag{9}$$
$$\boldsymbol{\beta}_{\text{dummies}} \sim \mathcal{N}(0, 0.2) \tag{10}$$

Priors Priors were selected to reflect weakly informative beliefs consistent with known biological constraints and to regularize the estimation of growth parameters in the presence of covariates. The asymptotic length L_∞ was assigned a normal prior centered at 500 mm, with a data-informed standard deviation drawn from a half-normal distribution $\mathcal{HN}(2)$. This reflects prior knowledge that asymptotic sizes for rainbow trout in similar systems typically fall between 400–600 mm, while still allowing flexibility. The observation error σ_{obs} and the hyperparameter σ_{L_∞} were each given half-normal priors with scale 2, reflecting reasonable uncertainty around measurement noise and biological variability. The intercept and continuous covariate effects on the growth coefficient k were modeled with standard normal priors $\mathcal{N}(0, 1)$, supporting broad exploration of potential influences while enforcing shrinkage toward zero in the absence of signal. Seasonal dummy effects were assigned tighter priors $\mathcal{N}(0, 0.2)$ to prevent overfitting from sparse or collinear binary predictors and to reflect their role as adjustment factors rather than primary drivers. All priors were selected to ensure identifiability and facilitate convergence during sampling, without imposing strong constraints on parameter estimates. The model was implemented in *NumPyro* [30], and posterior samples were obtained via Markov Chain Monte Carlo (MCMC) using the No-U-Turn Sampler (NUTS) [31].

Formulation The joint posterior distribution of the model parameters

$$\theta = \{L_\infty, \sigma_{\text{obs}}, \alpha_k, \boldsymbol{\beta}_k, \boldsymbol{\beta}_{\text{dummies}}\}$$

given the observed data

$$\mathcal{D} = \{L_1^{(i)}, L_2^{(i)}, t^{(i)}, \mathbf{X}_c^{(i)}, \mathbf{X}_d^{(i)}\}_{i=1}^N$$

is proportional to the product of the likelihood and prior distributions:

$$p(\theta \mid \mathcal{D}) \propto \prod_{i=1}^{N} \mathcal{N}\left(L_2^{(i)} \mid \hat{L}_2^{(i)},\ \sigma_{\text{obs}}^2\right) \times p(L_\infty) \times p(\sigma_{\text{obs}}) \times p(\alpha_k) \times p(\boldsymbol{\beta}_k) \times p(\boldsymbol{\beta}_{\text{dummies}}) \tag{11}$$

where the predicted length at recapture is defined as:

$$\hat{L}_2^{(i)} = L_\infty - (L_\infty - L_1^{(i)}) \cdot \exp\left(-k^{(i)} t^{(i)}\right) \tag{12}$$

and the individual-specific growth coefficient $k^{(i)}$ is modeled as:

$$k^{(i)} = \exp\left(\alpha_k + \mathbf{X}_c^{(i)} \boldsymbol{\beta}_k + \mathbf{X}_d^{(i)} \boldsymbol{\beta}_{\text{dummies}}\right) \tag{13}$$

Bayesian Gompertz Model Formulation

The second biological model implemented was the [14] growth equation, a sigmoid-shaped function that describes asymptotic growth using a double-exponential form. Originally developed to model human mortality, the Gompertz function has since been widely applied in biological and ecological contexts to capture growth processes that begin rapidly but decelerate over time as they approach an upper physiological limit. The model assumes that the relative growth rate decreases exponentially with time, making it particularly well-suited for species exhibiting rapid early development followed by gradual slowing as size approaches a maximum threshold. The Gompertz equation models the predicted fork length L_2 at time t as:

$$L_2 = L_\infty \cdot e^{-e^{-Kt}} \tag{14}$$

To estimate the parameters of the Gompertz growth model using Bayesian inference, we specified the following hierarchical model:

$$L_2 \sim \mathcal{N}(\hat{L}_2, \sigma_{\text{obs}}^2) \tag{15}$$

$$\hat{L}_2 = L_\infty \cdot \exp\left(-\exp(-k \cdot t)\right) \tag{16}$$

$$k = \exp\left(\alpha_k + \mathbf{X}_c \boldsymbol{\beta}_k + \mathbf{X}_d \boldsymbol{\beta}_{\text{dummies}}\right) \tag{17}$$

where:

- L_2: fork length at recapture,
- t: time at large (in years),
- L_∞: asymptotic maximum fork length,
- k: growth rate coefficient (positive, covariate-dependent),

- $\mathbf{X}_c$: matrix of scaled continuous covariates,
- $\mathbf{X}_d$: matrix of dummy variables for seasonal effects,
- $\boldsymbol{\beta}_k$: vector of regression weights for continuous covariates,
- $\boldsymbol{\beta}_{\text{dummies}}$: vector of regression weights for dummy predictors.

The prior distributions were specified as:

$$L_\infty \sim \mathcal{N}(500,\ \sigma^2_{L_\infty}), \quad \sigma_{L_\infty} \sim \mathcal{HN}(2) \tag{18}$$
$$\sigma_{\text{obs}} \sim \mathcal{HN}(2) \tag{19}$$
$$\alpha_k \sim \mathcal{N}(0, 1) \tag{20}$$
$$\boldsymbol{\beta}_k \sim \mathcal{N}(0, 1) \tag{21}$$
$$\boldsymbol{\beta}_{\text{dummies}} \sim \mathcal{N}(0, 0.2) \tag{22}$$

Priors Priors for the Gompertz model were specified identically to those in the Fabens VBGM to ensure consistency and comparability across formulations. All priors reflect weakly informative biological expectations while promoting identifiability. The asymptotic fork length L_∞ was assigned a normal prior centered at 500 mm, with a scale governed by a half-normal distribution $\mathcal{HN}(2)$, allowing flexibility while anchoring estimates within empirically observed trout size ranges. Both σ_{obs} and σ_{L_∞} were given half-normal priors to constrain positive variances without imposing rigid assumptions. The intercept α_k and continuous covariate coefficients $\boldsymbol{\beta}_k$ were assigned $\mathcal{N}(0, 1)$ priors to support moderate deviations while avoiding overdispersion. Dummy variable coefficients $\boldsymbol{\beta}_{\text{dummies}}$ were regularized with narrower $\mathcal{N}(0, 0.2)$ priors to mitigate overfitting from sparse or collinear seasonal predictors. By applying the same prior structure, we ensure that posterior differences between models are attributable to differences in functional form, rather than prior-induced biases. The model was also implemented in *NumPyro* [30], and posterior sampling was conducted using NUTS [31]).

Formulation The joint posterior distribution of the model parameters

$$\theta = \{L_\infty, \sigma_{\text{obs}}, \alpha_k, \boldsymbol{\beta}_k, \boldsymbol{\beta}_{\text{dummies}}\}$$

given the observed data

$$\mathcal{D} = \{L_2^{(i)}, t^{(i)}, \mathbf{X}_c^{(i)}, \mathbf{X}_d^{(i)}\}_{i=1}^N$$

is proportional to the product of the likelihood and prior distributions:

$$p(\theta \mid \mathcal{D}) \propto \prod_{i=1}^{N} \mathcal{N}\left(L_2^{(i)} \mid \hat{L}_2^{(i)},\ \sigma^2_{\text{obs}}\right) \times p(L_\infty) \times p(\sigma_{\text{obs}}) \times p(\alpha_k) \times p(\boldsymbol{\beta}_k) \times p(\boldsymbol{\beta}_{\text{dummies}}) \tag{23}$$

where the predicted length at recapture is defined as:

$$\hat{L}_2^{(i)} = L_\infty \cdot \exp\left(-\exp\left(-k^{(i)} t^{(i)}\right)\right) \tag{24}$$

and the individual-specific growth rate is:

$$k^{(i)} = \exp\left(\alpha_k + \mathbf{X}_c^{(i)} \boldsymbol{\beta}_k + \mathbf{X}_d^{(i)} \boldsymbol{\beta}_{\text{dummies}}\right) \tag{25}$$

Bayesian Justification.

We employed Bayesian inference for both the Fabens VBGM and Gompertz growth models due to its advantages in handling ecological data with complex structure and limited information. Bayesian methods allow for the incorporation of biologically informed prior knowledge–such as plausible bounds for asymptotic size or expected variability in growth rates–which improves parameter regularization and interpretability. These methods do require scaling of the non-indicator independent variables. This is especially critical in hierarchical models with latent variables and covariate-dependent growth rates, where frequentist estimators may yield unstable or biologically implausible results. Additionally, Bayesian models yield full posterior distributions, which enable more comprehensive uncertainty quantification and straightforward propagation of parameter uncertainty through model predictions. The Bayesian framework also supports posterior predictive checks, probabilistic forecasts, and model comparison using information criteria and cross validation tools that are not as readily applicable or interpretable under a frequentist paradigm. Taken together, these features make the Bayesian approach more robust, flexible, and ecologically appropriate for modeling rainbow trout growth under temporally and spatially structured environmental conditions.

Statistical and ML Models

In addition to the biologically grounded growth models, we implemented a diverse suite of statistical and ML algorithms to evaluate predictive performance across modeling paradigms. These included a Bayesian linear model for interpretable, uncertainty-aware regression; a Linear SVR model for capturing linear relationships with regularization [32]; a kernel-optimized SVR (with permutative importance), and several nonlinear, tree-based ensemble methods, including Random Forest (RF) [33], XGBoost [34], and LightGBM (LGBM) [27], each known for their robustness, embedded feature selection, and scalability. We also implemented an Artificial Neural Network (ANN) [35], representing a high-capacity function approximator capable of modeling complex, nonlinear relationships. Finally, to leverage the complementary strengths of mechanistic and data-driven approaches, we constructed two heterogeneous ensembles that combined the best-performing biological model with the top-performing ML model. This ensemble framework was designed to enhance predictive accuracy while preserving biological interpretability, aligning with recent advances in integrative ecological modeling.

Bayesian Linear Model Formulation

We implemented a Bayesian linear model to serve as a baseline for evaluating growth prediction under additive linear assumptions. The model assumes a normal likelihood with a linear predictor and constant variance:

$$y_i \sim \mathcal{N}(\mu_i, \sigma^2) \tag{26}$$

$$\mu_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} \tag{27}$$

where:

- y_i: fork length at recapture (response variable),
- x_{ij}: value of the jth predictor for observation i,

- β_0: model intercept,
- $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_p\}$: regression coefficients,
- σ: observation noise (residual standard deviation).

We placed weakly informative priors on all regression parameters to stabilize estimation without imposing strong constraints. The regression coefficients were assigned independent normal priors:

$$\beta_j \sim \mathcal{N}(0, 5), \quad \text{for } j = 0, 1, \ldots, p$$

and the standard deviation of residuals was parameterized as:

$$\sigma \sim \text{LogNormal}(0, 1)$$

with a deterministic transformation for interpretability:

$$\log(\sigma) \sim \mathcal{N}(0, 1)$$

We chose a Bayesian approach for the linear model to maintain consistency across other modeling frameworks and to take advantage of posterior inference, particularly under uncertainty and moderate sample sizes. Bayesian linear regression provides full posterior distributions for all parameters, allowing direct probabilistic interpretation and better uncertainty quantification than frequentist confidence intervals. Additionally, posterior draws for regression coefficients enable direct comparison with effect estimates from the nonlinear growth models. This formulation also improves robustness to multicollinearity and overfitting in high-dimensional predictor sets through regularization induced by priors. Unlike frequentist ordinary least squares (OLS), which provides only point estimates and asymptotic standard errors, the Bayesian version yields complete joint distributions and facilitates posterior predictive checks and model averaging. Moreover, by expressing the observation error in log-space, we achieved more stable sampling and better convergence behavior in MCMC, particularly under heteroscedastic conditions. The model was implemented in *NumPyro*, and posterior samples were drawn using NUTS [31].

Random Forest (RF)

An RF [33] is an ensemble learning method that constructs a collection of decorrelated decision trees and aggregates their predictions to produce a more stable and accurate model. For regression tasks, the predicted outcome $\hat{y}$ is computed as the average of T individual decision tree predictions:

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} f_t(\mathbf{x}) \tag{28}$$

where:

- T is the number of trees in the forest,
- $f_t(\mathbf{x})$ is the prediction of the t-th tree.

Each decision tree is trained on a bootstrap sample drawn with replacement from the original dataset, a process known as *bagging* (bootstrap aggregating). Furthermore, at each node split, RF consider only a random subset of predictors, introducing additional decorrelation between trees and reducing overfitting.

By averaging over many uncorrelated trees, RF significantly reduces the variance of individual tree estimators. While single decision trees are prone to high variance and may overfit training data, the ensemble average acts as a variance stabilizer, yielding a more generalizable model. This property is particularly advantageous in ecological datasets, which often contain complex interactions, nonlinearities, and noisy measurements. Each individual tree uses a greedy algorithm to recursively partition the feature space based on impurity measures such as Mean Squared Error (MSE) used in this study.

Importantly, RF offer a built-in mechanism for assessing feature importance. This importance is typically calculated (as it is here) based on the average decrease in impurity (or prediction error) resulting from splits on each feature, aggregated across all trees.

To optimize the performance of the Random Forest model, we employed a randomized hyperparameter search using 5-fold cross-validation. The search focused on key parameters that influence model complexity and generalization, including the number of trees, maximum depth of each tree, the minimum number of samples required to split an internal node, the minimum number of samples required to form a leaf, and the strategy used to select input features at each split. Rather than exhaustively evaluating all possible combinations, we sampled 25 random configurations from the defined hyperparameter space to balance efficiency and thoroughness. Model performance was assessed using negative mean squared error as the scoring metric, and the configuration yielding the lowest average validation error was selected as the final model.

The Random Forest model was tuned using a randomized search across the following hyperparameter space, with the selected values shown in **bold**:

- Number of trees (`n_estimators`): [100, **200**, 300, 400]
- Maximum tree depth (`max_depth`): [15, **25**, None]
- Minimum samples required to split a node (`min_samples_split`): [**2**, 5, 10]
- Minimum samples required to form a leaf node (`min_samples_leaf`): [**1**, 3, 5]
- Feature sampling method for tree construction (`max_features`): [**sqrt**, log2]

These settings were selected based on performance using randomized search with 5-fold cross-validation, optimizing for generalization and predictive accuracy.

XGBoost and LightGBM (LGBM)

Both XGBoost [34] and LGBM [27] are gradient-boosted decision tree (GBDT) frameworks that build powerful predictive models through the sequential addition of weak learners (typically shallow regression trees). The prediction function is constructed as a sum of T individual tree functions:

$$\hat{y} = \sum_{t=1}^{T} f_t(\mathbf{x}), \quad f_t \in \mathcal{F} \tag{29}$$

where $\mathcal{F}$ is the space of regression trees, and each f_t represents a base learner trained to predict the residuals of the current ensemble.

Unlike bagging-based methods like RF, boosting operates sequentially, where each new tree corrects the prediction error made by the ensemble so far. Both XGBoost and LGBM use gradient descent to minimize an objective function that combines a

differentiable loss function $\mathcal{L}(\hat{y}, y)$ (typically squared error for regression) with a regularization term $\Omega(f)$ to penalize model complexity:

$$\mathcal{L}_{\text{total}} = \sum_{i=1}^{n} \mathcal{L}(y_i, \hat{y}_i) + \sum_{t=1}^{T} \Omega(f_t) \tag{30}$$

Regularization discourages overfitting and promotes generalization by penalizing trees with excessive depth, number of leaves, or leaf weights.

XGBoost [26] constructs trees *depth-wise*, growing each level of the tree simultaneously and splitting all nodes at the current depth before proceeding. This strategy produces balanced trees and generally leads to faster convergence with fewer splits.

LGBM [27], in contrast, adopts a *leaf-wise* growth strategy: it selects the leaf with the largest reduction in loss and continues splitting from that point. This often results in deeper, more asymmetric trees, allowing for greater flexibility and improved computational efficiency on large datasets. However, it can increase the risk of overfitting on smaller datasets if not properly regularized.

To fine-tune the XGBoost model, we conducted a randomized hyperparameter search using 5-fold cross-validation. The search targeted parameters that directly influence model flexibility, learning dynamics, and regularization. These included the learning rate, maximum tree depth, number of boosting rounds, subsampling ratio for training instances, proportion of features used per tree, and both L_1 and L_2 regularization strengths. Given the use of higher learning rates, the number of boosting rounds was kept moderate to avoid overfitting. We sampled 25 random combinations from the defined hyperparameter space to ensure efficient yet thorough exploration. Model performance was evaluated based on negative mean squared error, and the configuration yielding the lowest average validation error was selected as the final model.

The XGBoost model was tuned using randomized search with 5-fold CV over the following hyperparameter space. The selected values are shown in **bold**:

- Learning rate (`eta`): [**0.05**, 0.1, 0.15, 0.2]
- Maximum tree depth (`max_depth`): [3, **4**, 5]
- Number of boosting rounds (`n_estimators`): [**200**, 400, 600, 800]
- Subsample ratio of training instances (`subsample`): [0.7, **0.8**, 0.9]
- Proportion of features used per tree (`colsample_bytree`): [0.8, 0.9, **1.0**]
- L1 regularization term on weights (`reg_alpha`): [**0**, 0.1]
- L2 regularization term on weights (`reg_lambda`): [**1**, 1.5]

These parameters were selected to balance generalization and computational efficiency, with model performance validated using cross-validation.

For the LightGBM model, we also implemented a randomized hyperparameter search using 5-fold cross-validation. The search was designed to explore parameters central to boosting performance and model regularization. These included the number of boosting rounds, learning rate, maximum tree depth, and the number of leaves per tree, which is particularly influential in LightGBM's gradient-based framework. Additional parameters such as subsampling ratios for rows and columns, along with L_1 and L_2 regularization strengths, were also tuned to improve generalization. A total of 20 random configurations were sampled from the defined hyperparameter space to achieve an efficient and targeted search. Model performance was evaluated using

negative mean squared error, and the configuration with the lowest average validation error was selected as the final model.

The LightGBM model was tuned using randomized search across the following hyperparameter space. Selected values are shown in **bold**:

- Number of boosting rounds (`n_estimators`): [100, **150**, 200]
- Learning rate (`learning_rate`): [**0.05**, 0.1, 0.15]
- Maximum tree depth (`max_depth`): [15, **20**, -1]
- Number of leaves per tree (`num_leaves`): [**31**, 50, 100]
- Subsample ratio of training data (`subsample`): [0.7, 0.8, **0.9**]
- Column sampling ratio per tree (`colsample_bytree`): [0.8, **0.9**, 1.0]
- L1 regularization term on weights (`reg_alpha`): [0, **0.1**]
- L2 regularization term on weights (`reg_lambda`): [**0**, 0.1, 1]

Support Vector Regression (SVR)

SVR extends the foundational concepts of Support Vector Machines (SVMs) to regression tasks by constructing an optimal predictive function that balances model complexity with tolerance for small prediction errors [32]. SVR estimates a continuous response variable $\hat{y}$ using a linear or nonlinear mapping:

$$\hat{y} = \mathbf{w}^\top \phi(\mathbf{x}) + b$$

where $\phi(\cdot)$ is a (possibly nonlinear) transformation of the input vector $\mathbf{x}$, $\mathbf{w}$ is the weight vector, and b is the bias term.

Unlike OLS, which minimizes squared error, SVR introduces an ε-insensitive loss function that ignores prediction errors within a margin of ε. The optimization objective is to find a "flat" function that approximates the data within this margin while minimizing complexity. Formally, the primal optimization problem follows.

$$\min_{\mathbf{w},b,\xi_i,\xi_i^*} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

subject to:

$$\begin{aligned} y_i - \mathbf{w}^\top\phi(\mathbf{x}_i) - b &\le \varepsilon + \xi_i \\ \mathbf{w}^\top\phi(\mathbf{x}_i) + b - y_i &\le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* &\ge 0 \end{aligned}$$

Here, $\|\mathbf{w}\|^2$ penalizes model complexity to promote generalization. The slack variables ξ_i and ξ_i^* capture deviations beyond the ε-tube, and the regularization parameter C governs the trade-off between model flatness and tolerance to errors exceeding ε.

This convex quadratic programming formulation is typically solved in the dual using Lagrangian multipliers, which naturally identifies a subset of *support vectors*–data points that lie outside the ε-tube and directly influence the regression function.

Solving the dual problem enables the use of *kernel functions*, defined as $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top\phi(\mathbf{x}_j)$, to compute inner products in high-dimensional feature spaces without explicit transformation. This *kernel trick* allows SVR to model nonlinear relationships efficiently. Common choices include the radial basis function (RBF), polynomial, and sigmoid kernels.

To develop an SVR with interpretable coefficients, we initially restricted the hyperparameter search to a linear kernel. This choice facilitates straightforward interpretation of the feature effects on the predicted outcome and allows comparison with traditional parametric models. While linear regression models also offer interpretable coefficients, we included the linear-kernel SVR model to assess whether a margin-based linear method with regularization could improve generalization performance without sacrificing interpretability. A randomized hyperparameter search with 5-fold cross-validation was employed to tune the penalty parameter, the epsilon-insensitive loss margin, and the kernel coefficient. Model performance was evaluated using negative root mean squared error (RMSE), and the configuration with the lowest average validation error was selected as the final model.

The linear Support Vector Regression (SVR) model was tuned (5-fold CV) using a predefined hyperparameter space. Since a linear kernel was used, the `gamma` parameter was not applicable. The hyperparameter search space and selected values are shown below, with selected values in **bold**:

- Penalty parameter (C): [0.1, 1, 10, **100**]
- Epsilon in the loss function (ϵ): [0.01, 0.1, **0.5**]

These parameters were selected based on performance using cross-validation, with the linear kernel fixed.

A second Support Vector Regression (SVR) model was formulated to evaluate the optimal kernel type as well as the other parameters. The investigated components mirrored those of the linear SVR, with the addition of kernel selection. Three kernels (linear, polynomial, and radial basis function or RBF) were examined using hyperparameter tuning and 5-fold cross-validation. A nonlinear Support Vector Regression (SVR) model was tuned to evaluate kernel performance, including linear, polynomial, and radial basis function (RBF) kernels. The selected configuration used an **RBF kernel**, with other hyperparameters optimized via randomized search and 5-fold cross-validation. The hyperparameter space and selected values are listed below, with selected values in **bold**:

- Kernel type (`kernel`): [linear, poly, **rbf**]
- Penalty parameter (C): [0.1, 1, 10, **100**]
- Epsilon in the loss function (ϵ): [**0.01**, 0.1, 0.5]
- Gamma (γ): [**scale**, 0.01, 0.1]

This configuration was found to provide the best trade-off between model complexity and predictive performance on the validation folds. To compare the parameter relevance of the SVR model against other ML approaches, **permutative feature importance** was evaluated. This method allowed for a consistent framework to assess the relative contribution of each input variable across models, enabling a robust comparison of feature influence in the presence of nonlinearities and interactions.

Artificial Neural Network (ANN)

Artificial Neural Networks (ANNs) [35] are layered computational models inspired by the architecture of the human brain. They are capable of approximating highly nonlinear, hierarchical relationships between inputs and outputs through learned feature transformations. In their most basic form, ANNs consist of an input layer, one or more hidden layers, and an output layer, with each layer composed of nodes (neurons) that

compute weighted combinations of their inputs. Although artificial neural networks (ANNs) are not always the top-performing choice for structured tabular data, especially when sample sizes are moderate, their inclusion in this study was motivated by both methodological and ecological considerations. From a methodological perspective, ANNs serve as a flexible, universal function approximator capable of capturing complex, nonlinear relationships that may elude simpler parametric models. Given the known nonlinearities in fish growth and environmental interactions, this capacity warranted their inclusion. Ecologically, neural networks allow us to test whether non-tree-based nonlinear models can uncover complementary patterns in rainbow trout growth dynamics, especially in interaction-heavy scenarios where traditional models may oversimplify relationships. Including the ANN also provided a useful baseline for comparing deep learning performance against more interpretable models such as tree ensembles and biologically grounded equations.

The general forward pass for an L-layer feedforward neural network is expressed as:

$$\hat{y} = \sigma^{(L)}\Bigg(W^{(L)} \cdot \sigma^{(L-1)}\Bigg(W^{(L-1)} \cdot \ldots \sigma^{(1)}\big(W^{(1)}\mathbf{x} + \mathbf{b}^{(1)}\big) + \cdots + \mathbf{b}^{(L-1)}\Bigg)\Bigg) + \mathbf{b}^{(L)} \tag{31}$$

where:

- $W^{(l)}$ denotes the weight matrix at layer l,
- $\mathbf{b}^{(l)}$ is the bias vector at layer l,
- $\sigma^{(l)}$ is the nonlinear activation function at layer l, such as the ReLU, sigmoid, or tanh function.

ANNs are trained by minimizing a loss function–such as mean squared error in regression–via backpropagation and stochastic gradient descent (SGD) or adaptive optimization algorithms like Adam. During training, weights and biases are iteratively adjusted to reduce the prediction error. The flexibility of ANNs comes at the cost of increased computational demand and sensitivity to hyperparameter choices, such as learning rate, batch size, and architecture depth.

The universal approximation theorem guarantees that a sufficiently large neural network can approximate any continuous function on a compact domain to arbitrary accuracy. This makes ANNs powerful tools for modeling nonlinear ecological systems with high-dimensional interactions that may be difficult to capture with traditional parametric models.

To identify an optimal configuration for the neural network model, we performed a randomized hyperparameter search using 5-fold cross-validation. The search space included a range of architectural and training parameters, such as the number and size of hidden layers, dropout rates, learning rates, weight decay values, learning rate scheduling parameters (initial cycle length and minimum learning rate), and batch sizes. From a total of 648 possible configurations, five hyperparameter combinations were randomly sampled and evaluated. For each combination, five models were trained and validated across cross-validation folds, with mean squared error used as the evaluation metric. The average validation loss across 5-folds was used to determine the most effective configuration. This initial tuning phase provided a computationally efficient yet rigorous method for selecting a high-performing neural network architecture consistent with the evaluation procedures used for other models in the study.

The best-performing artificial neural network (ANN) configuration was selected based on average validation loss across five cross-validation folds. The selected hyperparameters follow.

- Hidden layer sizes (`hidden_sizes`): [**256**, **128**, **64**]
- Dropout rates per layer (`dropout_rates`): [**0.4**, **0.3**, **0.2**]
- Learning rate (`learning_rate`): **0.0005**
- Weight decay (`weight_decay`): **5e-5**
- Learning rate restart interval (`lr_restart_interval`): **20**
- Minimum learning rate (`min_lr`): **1e-6**
- Batch size (`batch_size`): **32**

Results

The comparative analysis revealed substantial performance differences among biological and ML models in predicting rainbow trout growth. The comparative analysis revealed meaningful differences in predictive accuracy across mechanistic, Bayesian, and machine-learning approaches, while also showing that several top-performing models produce near-equivalent error on the held-out test set. We therefore emphasize not only point performance (RMSE/MAE), but also interpretability, model parsimony, and probabilistic dominance relationships across the full posterior error distribution.

Descriptive Statistics

Table 1 presents the key descriptive statistics for several numerical variables used in the analysis of Rainbow Trout growth and migration. The "Time at Large" variable, with a mean of 243.47 days and a standard deviation of 285.59, displays substantial variability, further confirmed by its high skewness (2.43) and kurtosis (6.92), indicating the presence of long-tailed observations and potential outliers. The physical growth measures, such as L_1 and L_2, show relatively symmetric distributions with low skewness and slightly negative kurtosis, suggesting a mild platykurtic shape. Environmental variables such as "Water Temperature" and "Solar Insolation" exhibit moderate variability and near-normal distributions, which may enhance model stability. Interestingly, "Release River Mile" is negatively skewed with an extremely high kurtosis (13.87), suggesting a heavily peaked distribution with frequent values near one end.

Table 1. Descriptive statistics of selected numerical variables

Variable	Mean	Std	Min	Median	Max	Skew	Kurtosis
Release River Mile	-3.74	0.69	-14.89	-3.75	-1.96	-1.12	13.87
Time at Large	243.47	285.59	28.00	130.50	2256.00	2.43	6.92
Fork Length at Release (L_1)	187.26	87.53	70.00	163.00	457.00	0.36	-1.24
Fork Length at Recapture (L_2)	218.92	85.79	74.00	230.00	472.00	0.06	-1.30
Weight at Release	118.78	132.94	3.50	51.00	1182.00	1.39	1.98
Discharge	18.65	4.14	11.33	17.81	32.11	1.34	1.74
Water Temperature	16.86	3.89	8.65	16.05	27.14	0.61	0.01
Solar Insolation	50.24	17.86	19.18	49.71	105.92	0.50	0.09
Soluble Reactive P Concentration	0.01	0.00	0.00	0.01	0.02	0.42	-0.27
Rainbow Trout Biomass (kg)	12.38	6.09	5.56	10.45	34.03	0.99	0.20

Baseline Models

The baseline VBGM and Gompertz function were evaluated as mechanistic benchmarks, without the inclusion of covariates. The results indicate a clear performance difference between the two models. The baseline Gompertz model achieved a lower RMSE (61.52 mm) and mean absolute error (MAE = 49.81 mm) compared to the baseline VBGM (RMSE = 86.29 mm, MAE = 77.19 mm), suggesting superior predictive accuracy. Additionally, the Gompertz model explained a substantially higher proportion of variance in the test data, with an R^2 of 0.492 and adjusted R^2 of 0.483, whereas the VBGM performed no better than a mean-only model ($R^2 = 0.000$).
Information-theoretic criteria also favored the Gompertz formulation, which yielded substantially lower Akaike Information Criterion (AIC=32658.82), corrected AIC (32660.38), and Bayesian Information Criterion (BIC=32940.17) values than the VBGM (AIC = 34648.67, AICc = 34650.23, BIC = 34930.02). These results support the Gompertz model as the more effective baseline structure for capturing growth patterns in this dataset. Again, these models served as the comparative baseline.

Bayesian Models

The first three non-baseline model results were all Bayesian. In all cases, the convergence diagnostics indicated excellent mixing and stability across all parameters. In all cases, estimation was conducted using Markov Chain Monte Carlo (MCMC) with 4 chains and 1500 draws per chain (500 tuning, 1,000 posterior). The target accept=0.95 parameter increased the robustness of the NUTS in exploring complex posteriors.

The Gelman–Rubin statistic ($\hat{R}$) was 1.00 for all parameters across all models, demonstrating full convergence. Effective sample sizes (ESS) for both the bulk and tail of the posterior distributions were well above the commonly recommended threshold of 400, with many exceeding 6,000, indicating that the chains produced a large number of effectively independent samples. Moreover, no divergences or energy transition issues were reported for any of the models, further confirming their stability and reliability of the model.

Fabens VBGM

The VBGM assumes that growth decelerates asymptotically as an organism approaches its theoretical maximum size L_∞. The key strength of this model is that it provides posterior distributions over all unknowns, including individual-level growth rates, allowing for full uncertainty quantification. Unlike traditional VBGM approaches that assume constant growth rates, this model incorporates individual heterogeneity through environmental and temporal covariates, which is critical in ecological modeling where growth is often affected by dynamic habitat conditions.

The performance metrics for the VBGM suggest a solid, biologically informed baseline. The model yielded an RMSE of 16.82 and an MAE of 11.60, indicating that the average deviation of predicted fish lengths from observed values was around 11 to 17 mm. From a biological standpoint, this level of error is relatively modest given the natural variability in growth patterns among fish populations, and it suggests that the model is effectively capturing the underlying growth dynamics. The R^2 value of 0.9620 and adjusted R^2 of 0.9619 imply that over 96% of the variability in observed lengths is explained by the model, reinforcing the biological plausibility of VBGM in modeling somatic growth processes. Additionally, the AIC (24,945.04), AICc (24,945.05), and BIC (24,962.99) metrics indicate strong model parsimony, supporting the idea that the growth curve is both statistically and biologically well-specified without unnecessary complexity. (While a Widely-Applicable Information Criterion (WAIC) is generally

preferred, use of the AIC, AICc, and BIC supported comparison.) Overall, these results affirm the VBGM's capacity to capture length-at-age patterns with high fidelity, aligning with theoretical expectations of metabolic scaling and resource allocation in fish development.

Gompertz Model

The Gompertz model also demonstrate reasonable but somewhat inferior predictive performance relative to the VBGM. The performance on the test set follows: RMSE=16.924, MAE=11.674, R^2=0.962, Adjusted R^2=0.962, AIC=24982.28, AICc=24982.29, BIC=25000.24.

While the Gompertz model outperformed the VGBM in the baseline, no-covariate setting likely due to its flexible curvature and early growth deceleration, its performance declined relative to VBGM once covariates were introduced in a Bayesian framework. One reason for this reversal is structural: the Gompertz model includes a single global rate parameter, which can limit its ability to accommodate feature-dependent variation in growth. In contrast, the VBGM, though initially more rigid, benefits more from the inclusion of biologically relevant covariates that can adjust parameters such as the asymptotic size or growth rate across individuals or conditions. This flexibility, combined with the regularizing effects of Bayesian estimation, enables the VBGM to better capture context-specific growth dynamics and reduces overfitting. The Gompertz model, by contrast, may become overparameterized or less stable when extended with covariates, diminishing its initial advantage.

Bayesian Linear Model

We fit a Bayesian linear regression model using *NumPyro* to estimate the relationship between trout recapture length (L_2) and a combination of continuous and categorical predictors. All continuous predictors were standardized prior to model fitting, while the response variable (L_2) remained in its original unit of millimeters to preserve ecological interpretability.

The model used a Normal prior for the coefficients ($\beta \sim \mathcal{N}(0, 5)$), a LogNormal prior on the standard deviation ($\log \sigma \sim \mathcal{N}(0, 1)$), and was sampled using the NUTS with 4 chains, each producing 1,500 posterior samples after a 500-sample tuning phase.

The Bayesian linear model demonstrated strong predictive accuracy compared to baseline methods. It achieved an R^2 of 0.911, with an adjusted R^2 of 0.910, indicating that the model captured the majority of the variance in the response variable. The RMSE was 25.72 mm, and the MAE was 18.54 mm, both reflecting substantial improvements over prior versions and competitive performance relative to more complex models.

The information criteria supported the model's improved fit: AIC = 27,533.81, AICc = 27,535.44, and BIC = 27,821.15. These values are substantially lower than those observed in earlier specifications and within a competitive range compared to tree-based or boosting models, suggesting that the revised Bayesian linear model more effectively captured the underlying structure of the data.

While the Bayesian framework offers clear benefits in interpretability and uncertainty quantification, its application here did not result in a competitive model. Future improvements may involve incorporating hierarchical structures, interaction terms, or more flexible priors to better accommodate ecological complexity in trout growth modeling.

Trees Models

The results for each of the tree models follow. While traditional information criteria such as AIC, AICC, and BIC are typically derived from maximum likelihood estimation (MLE), many modern ML models (such as tree-based ensembles) do not possess closed-form likelihoods. In these cases, we approximate information criteria by assuming Gaussian residuals and computing the residual sum of squares (RSS) to estimate a pseudo-likelihood. Specifically, we treat the residuals as arising from a normal distribution with constant variance, and derive a log-likelihood under this assumption. The number of effective model parameters is approximated by counting the trainable parameters (e.g., via PyTorch's `model.parameters()`), serving as a proxy for model complexity. While this approach does not provide true likelihood-based inference, it enables comparison across a diverse set of models using a consistent penalized error framework. As such, these criteria should be interpreted heuristically rather than as strict measures of statistical fit.

RF

The tuned Random Forest model was then applied to the pristine test set data, achieving an RMSE of 18.40 mm and a MAE of 12.20 mm. The model explained approximately 95.45% of the variance in the outcome ($R^2 = 0.9545$, adjusted $R^2 = 0.9538$), indicating excellent predictive accuracy. The information criteria values (AIC = 25,563.22, AICc = 25,564.78, and BIC = 25,844.57) further confirm the model's strong fit relative to other approaches. The three most important predictors identified by the Random Forest were L_1 (length at release), time at large, and weight at release. These are intuitive and biologically plausible variables, reinforcing their relevance in predicting growth outcomes.

XGBoost

XGBoost exhibited the lowest predictive error with an RMSE of 16.14 mm, slightly ahead of Random Forest (16.46 mm), VBGM (16.82 mm), and the Gompertz model (17.07 mm). Although the differences in RMSE are consistent across repeated validation folds, the absolute gap between the best-performing ML models and the Bayesian VBGM is under 1 mm. This suggests that the improvements in predictive accuracy, while statistically detectable, may not translate to biologically meaningful differences in growth estimation for management purposes.

LightGBM

LightGBM achieved an RMSE of 16.25 mm and an MAE of 10.80 mm, delivering predictive performance nearly equivalent to XGBoost. With an R^2 of 0.9645 and an adjusted R^2 of 0.9640, the model explained a similarly high proportion of variance in the outcome. Although slightly higher than XGBoost, the information criteria (AIC = 24,830.76, AICc = 24,832.32, and BIC = 25,112.11) still reflect a strong model fit and reinforce LightGBM's competitiveness among gradient boosting approaches. The top three most influential features were L_1 (length at release), time at large, and the release river map, indicating the model's ability to integrate both biological and spatial predictors effectively.

SVR Models

Linear SVR

The linear SVR model showed noticeably weaker performance compared to the top ensemble methods in this study. It yielded an RMSE of 25.35 mm and a MAE of 15.40 mm–substantially higher than those of XGBoost (RMSE = 16.14) and LightGBM (RMSE = 16.25), and even lagging behind the Random Forest model (RMSE = 18.40). With an R^2 of 0.9137 and adjusted R^2 of 0.9122, the SVR explained just over 91% of the variance in recapture length–respectable, but clearly lower than the approximately 96% explained by the best-performing models.

From a model selection perspective, SVR also fared less favorably. Its AIC = 27,446.61 and BIC = 27,727.96 were substantially higher than those for XGBoost (AIC = 24,790.38, BIC = 25,071.73) and LightGBM (AIC = 24,830.76, BIC = 25,112.11), indicating poorer balance between model fit and complexity. It also trailed the Random Forest in information criteria, suggesting that SVR was less efficient overall in modeling the observed data.

While the linear Support Vector Regression (SVR) model did not achieve top-tier predictive performance compared to nonlinear models, it offered a distinct advantage in interpretability. Specifically, linear SVR provides direct coefficient estimates, enabling a clear understanding of the direction and relative magnitude of each input variable's contribution to the predicted outcome. This directional insight is valuable for interpreting ecological or biological mechanisms in fish growth modeling. However, this transparency comes at the cost of flexibility, as the linear kernel assumes strictly additive and linear relationships between predictors and the response variable. Consequently, the model is unable to capture complex interactions or nonlinear effects, which likely explains its higher error metrics relative to kernel-based and ensemble models.

Kernel-Optimized SVR

The RBF-kernel SVR model demonstrated strong predictive performance, outperforming several baseline and traditional models. It achieved a RMSE of 18.81, an MAE of 11.77, and an R^2 value of 0.952. The adjusted R^2 was similarly high at 0.951, indicating robust model fit even after accounting for model complexity. In terms of information-theoretic criteria, the RBF SVR produced an AIC of 25692.93, AICC of 25694.49, and BIC of 25974.28. While not the top-performing model across all metrics, it offered a balance between accuracy and generalization. Notably, it outperformed the linear SVR and Bayesian linear models across every metric, demonstrating the benefit of modeling nonlinear relationships in fish growth data through kernel-based methods.

To evaluate the relative influence of input features in the RBF SVR model, we employed permutation feature importance, a model-agnostic approach that quantifies importance based on changes in prediction error. This method involves systematically shuffling the values of each feature in the test set and measuring the resulting increase in the model's prediction error, thereby capturing how much each feature contributes to the model's performance. Results are incorporated to the feature importances analyses.

ANN

The Artificial Neural Network (ANN) model (see Figure 2) delivered solid predictive performance, with an RMSE of 18.47 mm and a MAE of 13.09 mm. These results placed it within the same general performance range as the Random Forest (RMSE = 18.40), though notably behind gradient-boosted models such as XGBoost (RMSE = 16.14) and LightGBM (RMSE = 16.25). The network achieved an R^2 of 0.954 and an

adjusted R^2 of 0.952, indicating that it explained over 95% of the variance in recaptured length–strong, though slightly less than the top-performing boosting algorithms. While the ANN captured a meaningful portion of the underlying growth signal, its higher AIC (25,678.64), AICc (25,684.92), and BIC (26,241.34) suggest a trade-off in model efficiency compared to more parsimonious ensemble methods.

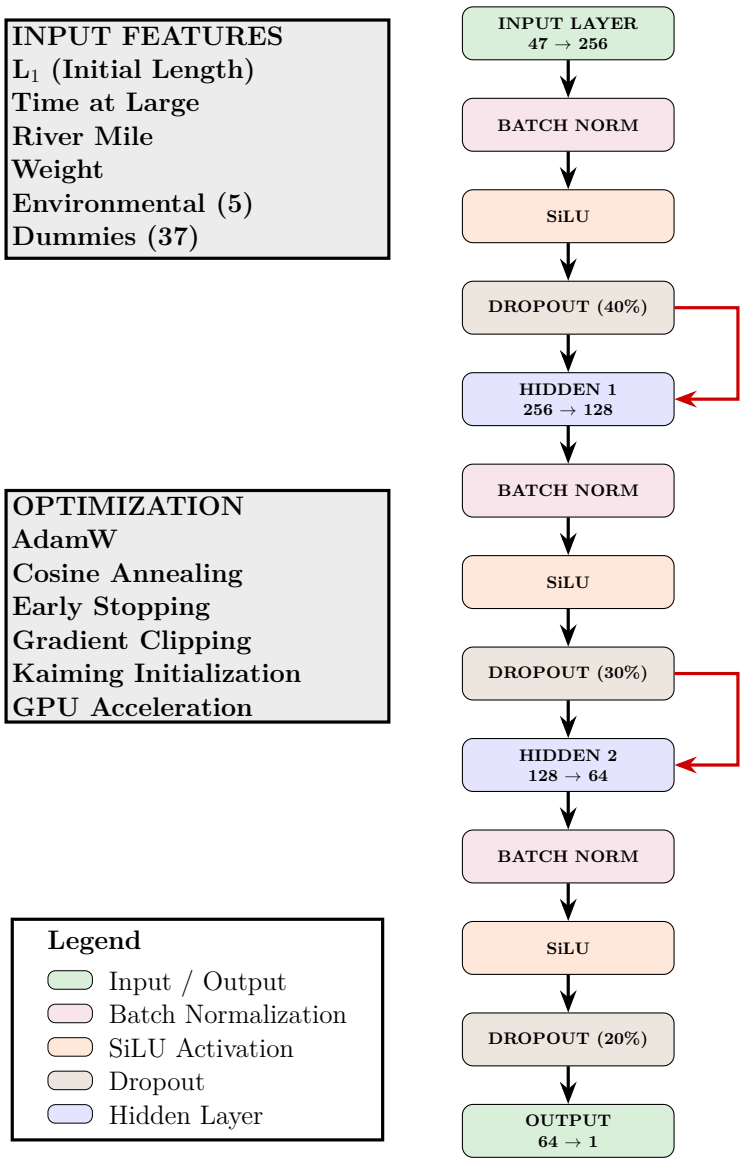

Fig 2. Neural network architecture with three hidden layers (256, 128, 64 neurons), batch normalization, SiLU activations, dropout regularization, and residual connections. The model maps 47 biological and environmental input features to a continuous prediction of fork length at recapture.

From an information-theoretic standpoint, the ANN yielded an AIC of 25,678.64 and a BIC of 26,241.34, both notably higher than those of the ensemble methods: XGBoost (AIC = 24,790.38, BIC = 25,071.73), LightGBM (AIC = 24,830.76, BIC = 25,112.11), and Random Forest (AIC = 25,563.22, BIC = 25,844.57). These elevated values suggest a more pronounced penalty for model complexity, which is consistent with expectations

given the parameter-rich architecture of neural networks. Unless effectively regularized, ANNs tend to be less parsimonious than tree-based methods. Integrated Gradients (IG), an attribution method that quantifies feature importance by integrating the gradients of the model's output with respect to its inputs along a straight-line path from a baseline (e.g., a zero vector) to the actual input, offer comparative insight into the internal prioritization of predictive factors by the network. IG are used for importance comparisons subsequently.

Ensembles

To synthesize the complementary strengths of biological and ML models, we implemented a stacked ensemble that combined predictions from the best-performing models within each paradigm as well as Bayesian Model Averaging (BMA). An explication follows.

For the first model, we used a linear regression model as a meta-learner to optimally weight predictions from the Bayesian VBGM and XGBoost. The meta-model was trained exclusively on out-of-sample predictions generated through cross-validation, ensuring that final ensemble weights were not influenced by data leakage. This stacking approach allowed the ensemble to adaptively balance interpretability and predictive flexibility, favoring one model or the other depending on local patterns in the data. By leveraging the structural insights of the biological model and the nonlinear predictive power of the ML model, the stacked ensemble achieved superior performance relative to all individual models. Its root mean square error (RMSE) was the lowest of any tested configuration, underscoring the practical value of integrating mechanistic understanding with data-driven learning in ecological forecasting.

The stacked ensemble, which combined predictions from the Bayesian VBGM and XGBoost, yielded the best overall performance across all evaluated models. It achieved the lowest RMSE (15.96 mm) and MAE (10.62 mm), indicating improved predictive accuracy over both its constituent models. The stacked ensemble also recorded the highest coefficient of determination ($R^2 = 0.9658$) and adjusted $R^2 = 0.9658$, reflecting its superior ability to explain variance in fork length growth while accounting for model complexity. From an information-theoretic perspective, it produced the lowest values for the Akaike Information Criterion (AIC = 24,635.91), corrected AIC (AIC_C = 24,635.92), and Bayesian Information Criterion (BIC = 24,653.87), further supporting its parsimony and goodness-of-fit relative to alternatives. These results confirm the effectiveness of the stacking strategy in integrating mechanistic and ML models, ultimately offering more accurate and robust growth forecasts than any single model alone.

The second approach, BMA, is a principled approach to model combination that accounts for uncertainty across competing models by averaging their predictions weighted by their posterior probabilities. Unlike traditional ensemble methods that may use equal or heuristic weights, BMA incorporates Bayesian reasoning to assign greater influence to models that explain the data well while penalizing overly complex or poorly fitting models. This results in predictive distributions that naturally reflect model uncertainty and often lead to improved generalization performance.

The results for the Bayesian Model Averaging (BMA) ensemble reveal its strength as a competitive and balanced forecasting approach. In terms of RMSE and MAE, BMA matches the performance of XGBoost exactly—achieving an RMSE of 16.137 and an MAE of 10.722—indicating that its averaged predictions closely align with those of the strongest individual model. However, BMA information criteria values are uniformly lower than or comparable to other models, notably outperforming stacked ensembles, which although slightly better in RMSE, incur higher complexity penalties. This reflects BMA's Bayesian advantage: it combines predictive strength with parsimony, minimizing

overfitting risk by integrating model uncertainty. Overall, BMA delivers a robust, interpretable alternative to black-box models by leveraging posterior-weighted synthesis while maintaining statistical efficiency.

Model Comparison

Model comparisons across metrics and coefficients identify both diverging and converging models. Explication follows.

Metrics

Table 2 summarizes the performance of all models using RMSE, MAE, R^2, adjusted R^2, and information criteria (AIC, AICC, BIC). Importantly, the leading models are tightly clustered in absolute error (typically within ≈ 1 mm RMSE of one another), so the practical distinction among top performers is not accuracy alone, but whether the method provides biological interpretability and uncertainty quantification alongside predictive power.

The best overall performance was achieved by the *Stacked Ensemble*, which outperformed all individual models with the lowest RMSE (15.96 mm), lowest MAE (10.62 mm), and highest R^2 (0.9658). This highlights the effectiveness of aggregating predictive strengths from both mechanistic and ML paradigms.

Among individual learners, *XGBoost* and *LightGBM* performed exceptionally well, achieving RMSE values of 16.14 mm and 16.25 mm, respectively, with high R^2 values and favorable information criteria. These results reaffirm the ability of gradient boosting algorithms to model complex nonlinear processes in ecological systems. Bayesian biological models, including the *VBGM* and the *Gompertz equation*, also ranked competitively, demonstrating that biologically grounded models remain viable, especially when well-calibrated, informed by data, and inclusive of covariates.

In contrast, the *Baseline VBGM* and *Baseline Gompertz* models—fitted without covariates—performed poorly, with RMSE values of 86.29 mm and 61.52 mm, respectively, and substantially lower R^2 values. These results emphasize the limitations of using biologically motivated models without accounting for relevant environmental and contextual variables, reinforcing the importance of covariate inclusion for predictive accuracy.

Outside of the poor-performing baseline models, the *Bayesian Linear Model* and *Linear SVR* showed the weakest performance among fully trained models, with RMSEs exceeding 25 mm and comparatively lower R^2 values, indicating a poor fit to the data. However, the *Radial Basis Function (RBF) SVR* improved substantially upon the linear variant, achieving an RMSE of 18.81 mm and $R^2 = 0.952$, illustrating the benefit of nonlinear kernel methods in capturing more complex relationships. The *Random Forest* and *Neural Network* models also produced R^2 values above 0.95, but incurred higher information criteria, reflecting greater model complexity and possibly reduced generalization.

These findings support a natural next step: applying *Bayesian Model Averaging (BMA)* to integrate top-performing models across paradigms. The BMA approach achieved similar accuracy to XGBoost while explicitly incorporating model uncertainty through posterior-weighted predictions. This technique may be particularly advantageous in ecological forecasting contexts where interpretability, uncertainty quantification, and generalizability are critical. Future research could assess whether BMA further enhances model robustness while preserving the domain-specific strengths of both mechanistic and ML-based approaches.

Table 2. Comparison of model performance metrics (sorted by RMSE). Baseline for percent reduction is the baseline von Bertalanffy growth model (VBGM). Abbreviations: VBGM = von Bertalanffy growth model; BMA = Bayesian Model Averaging; ANN = Artificial Neural Network; SVR = Support Vector Regression; RF = Random Forest; LGBM = Light Gradient Boosting Machine; XGB = XGBoost; RMSE = root mean squared error.

Model	RMSE	MAE	R^2	Adj R^2	AIC	AICC	BIC	% RMSE < VBGM
Ensemble (Stacked)	15.956	10.619	0.9658	0.9658	24635.91	24635.92	24653.87	81.5
XGBoost	16.137	10.722	0.9650	0.9644	24790.38	24791.94	25071.73	81.3
Ensemble (BMA)	16.137	10.722	0.9650	0.9650	24702.38	24702.39	24720.34	81.3
LightGBM	16.248	10.800	0.9645	0.9640	24830.76	24832.32	25112.11	81.2
Bayesian VBGM	16.817	11.600	0.9620	0.9620	24945.04	24945.05	24963.00	80.5
Bayesian Gompertz	16.924	11.674	0.9615	0.9615	24982.28	24982.29	25000.24	80.4
Random Forest	18.404	12.203	0.9545	0.9538	25563.22	25564.78	25844.57	78.7
ANN	18.471	13.089	0.9542	0.9526	25678.64	25684.92	26241.34	78.6
SVR RBF	18.814	11.772	0.9524	0.9517	25692.93	25694.49	25974.28	78.2
SVR Linear	25.352	15.395	0.9137	0.9122	27446.61	27448.17	27727.96	70.6
Bayesian Linear Model	25.723	18.544	0.9111	0.9096	27533.89	27535.52	27821.23	70.2
Baseline Gompertz	61.516	49.810	0.4916	0.4833	32658.82	32660.38	32940.17	28.7
Baseline VBGM	86.289	77.193	-0.0003	-0.0166	34648.67	34650.23	34930.02	0.0

Coefficients

Figure 3 illustrates the top nine continuous features identified through average normalized importance scores across all models. The feature importance scores were normalized using min-max scaling to a [0,1] range to enable direct comparison across different models and metrics. The variable L_1, representing asymptotic length, emerged as the most influential predictor overall. Its importance was consistently high across nearly all modeling approaches, with particularly strong influence in the Bayesian Gompertz, Bayesian VBGM, and SVR RBF models. This is expected, as L_1 is a core parameter in both the VBGM and the Gompertz models, where it serves as a theoretical maximum size. These mechanistic models explicitly rely on this variable to define growth trajectories, which explains its critical importance.

The variable *Time at Large*, another essential parameter in traditional growth models, also ranked highly in terms of predictive contribution. In both the VBGM and Gompertz frameworks, *Time at Large* maps directly to the growth curve's temporal component, linking observed size back to age or exposure duration. Its strong influence in ML models (including LightGBM and Neural Networks) underscores its general predictive relevance, even outside mechanistic contexts.

Additional high-ranking features included *Weight at Release*, *Rainbow Trout Biomass*, and *Soluble Reactive Phosphorous Concentration*. These environmental or biological covariates were less central in traditional growth models but proved informative in ensemble methods like XGBoost and Random Forest. This suggests that data-driven models may capture interaction effects and context-dependent relationships that lie outside the scope of classical equations.

The consistency of L_1 and *Time at Large* across both mechanistic and ML approaches validates their foundational role in growth modeling. Meanwhile, the variable-specific strengths of different algorithms highlight the complementary nature of ensemble and domain-specific methods when predicting ecological phenomena.

Bayesian Probabilistic Comparison of RMSE

Pairwise posterior probabilities were computed from the joint posterior distribution over model RMSEs to estimate the probability that one model outperformed another, as

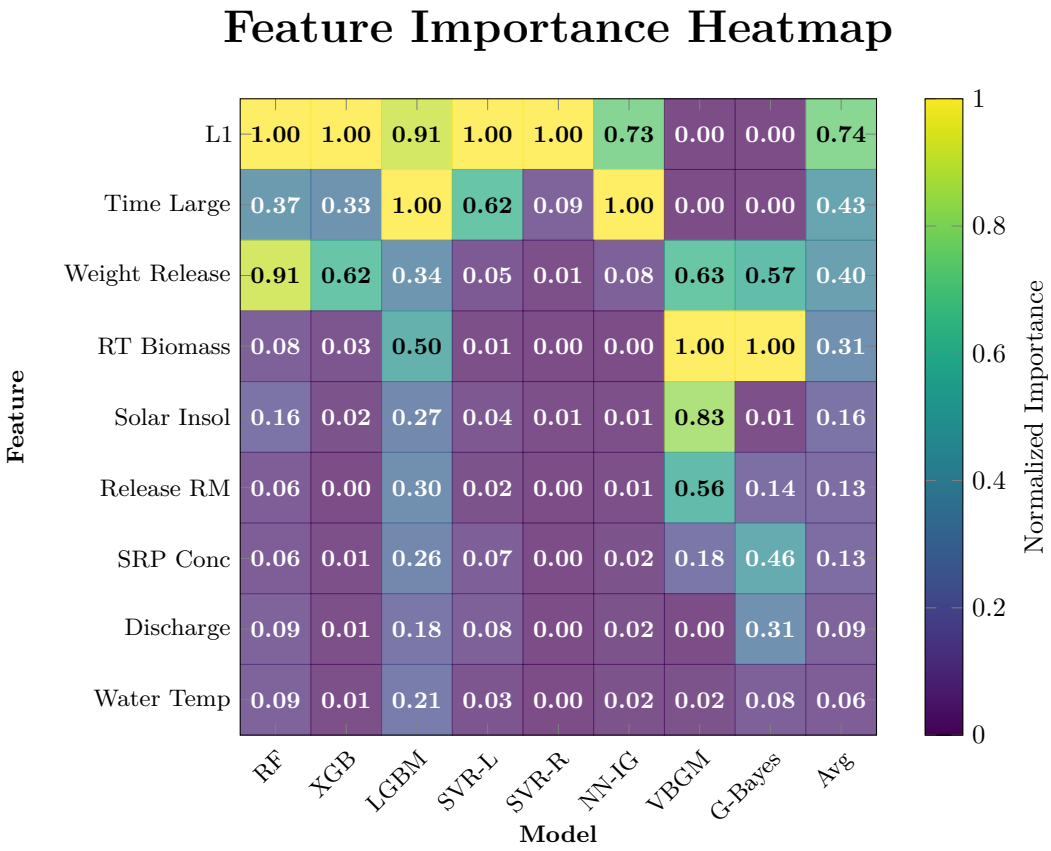

Fig 3. Feature importance heatmap showing normalized importance values (0-1) across nine models. Models: RF = Random Forest, XGB = XGBoost, LGBM = LightGBM, SVR-L = Support Vector Regression (Linear), SVR-R = Support Vector Regression (RBF), NN-IG = Neural Network with Integrated Gradients, VBGM = Von Bertalanffy Growth Model (Bayesian), G-Bayes = Gompertz (Bayesian), Avg = Average across all models. Features: L1 = Initial length, Time Large = Time at large, Weight Release = Weight at release, RT Biomass = Rainbow trout biomass, Solar Insol = Solar insolation, Release RM = Release river mile, SRP Conc = Soluble reactive phosphorous concentration, Water Temp = Water temperature. Color intensity and numerical values indicate relative feature importance within each model.

summarized in Table 3. These values represent the proportion of posterior samples for which the RMSE of one model was lower than that of another. The resulting dominance structure reveals a clear hierarchy, with the stacked ensemble stochastically dominating all individual models across the full error distribution, Bayesian Model Averaging and leading machine-learning methods forming a second tier, and the mechanistic VBGM variants, while substantially improved, remaining dominated when nonlinear environmental effects are present.

Table 3. Pairwise Stochastic Dominance Probabilities (Sorted by Number of Models Beaten)

Model	Ensemble (Stacked)	LightGBM	XGBoost	Ensemble (BMA)	Bayesian VBGM	Bayesian Gompertz	Random Forest	ANN	SVR_Radial	SVR_Linear	Bayesian Linear
Ensemble (Stacked)	-	0.512	0.509	0.506	0.525	0.535	0.589	0.596	0.612	0.771	0.776
LightGBM	0.488	-	0.501	0.501	0.524	0.529	0.585	0.583	0.598	0.770	0.775
XGBoost	0.491	0.499	-	0.504	0.527	0.543	0.583	0.593	0.598	0.765	0.775
Ensemble (BMA)	0.494	0.498	0.496	-	0.528	0.530	0.582	0.590	0.607	0.766	0.774
Bayesian VBGM	0.475	0.476	0.473	0.471	-	0.508	0.562	0.568	0.577	0.755	0.759
Bayesian Gompertz	0.465	0.471	0.457	0.470	0.491	-	0.550	0.569	0.580	0.756	0.758
Random Forest	0.411	0.415	0.417	0.418	0.438	0.450	-	0.503	0.520	0.724	0.727
ANN	0.404	0.417	0.407	0.410	0.432	0.431	0.497	-	0.519	0.722	0.724
SVR_Radial	0.388	0.402	0.402	0.393	0.423	0.420	0.480	0.480	-	0.715	0.726
SVR_Linear	0.229	0.230	0.235	0.234	0.245	0.244	0.276	0.279	0.285	-	0.514
Bayesian Linear	0.224	0.225	0.225	0.226	0.241	0.241	0.273	0.276	0.274	0.486	-

Discussion

Overall Synthesis

This study aimed to bridge traditional ecological modeling and modern machine learning (ML) methods to forecast invasive rainbow trout growth in the Lower Colorado River, highlighting the complementary strengths and trade-offs inherent in each modeling paradigm. Our findings are contextualized within existing literature, emphasizing both ecological theory and methodological advances.

The results reveal a clear structural pattern that goes beyond simple model ranking. While several modern machine–learning methods (XGBoost, LightGBM, Random Forest, and ANN) achieved strong predictive accuracy, the stacked ensemble consistently dominated across the full posterior distribution of errors, indicating not merely higher mean performance but greater reliability across ecological regimes. At the same time, the strong performance of covariate-augmented mechanistic models (VBGM and Gompertz) demonstrates that biological growth structure remains highly informative when environmental drivers are properly incorporated. Taken together, these findings show that trout growth in the Lower Colorado River is governed by both intrinsic biological constraints and complex nonlinear environmental interactions, and that hybrid ensemble frameworks are uniquely capable of capturing both.

The most striking result is the transformative impact of incorporating covariates and advanced modeling approaches. Relative to baseline models without covariates, the stacked ensemble reduced RMSE by 70.33 mm (81.5%) compared with the baseline VBGM and by 45.56 mm (74.1%) compared with the baseline Gompertz model. These large gains demonstrate the critical importance of environmental and biometric

covariates for biological growth modeling and highlight the severe limitations of traditional approaches that ignore ecological context.

The predictive results revealed important differences in model effectiveness. The stacked ensemble achieved the best point estimate performance (RMSE = 15.96 mm, R^2 = 0.9658) and exhibited the highest pairwise dominance probabilities across posterior RMSE distributions. This supports the effectiveness of hybrid approaches that integrate mechanistic structure with data-driven flexibility [18], showing that ensemble strategies can deliver both high average accuracy and consistent superiority across the full distribution of outcomes.

LightGBM (RMSE = 16.25 mm) and XGBoost (RMSE = 16.14 mm) formed a strong second tier of performance, with LightGBM achieving 9 dominance wins and XGBoost achieving 8 in the stochastic comparison. Both models produced over 80% RMSE reductions relative to baseline approaches, confirming the suitability of gradient boosting methods for ecological modeling as demonstrated by [15]. Our results extend this literature by showing that tree-based ensemble methods are particularly effective for biological growth prediction, while stacked ensembles provide additional gains through model integration.

Bayesian Model Averaging (BMA) also performed strongly, achieving an RMSE of 16.14 mm (R^2 = 0.9650) and 7 dominance wins in the stochastic analysis. By weighting models according to their posterior evidence, BMA provides a principled balance between predictive accuracy and explicit uncertainty quantification [16]. Relative to the baseline VBGM, BMA produced an 81.3% reduction in RMSE, demonstrating that probabilistic model averaging can deliver substantial performance gains while preserving inferential rigor.

Traditional mechanistic models remained valuable when enhanced with covariates and Bayesian estimation. The Bayesian VBGM achieved competitive predictive performance (RMSE = 16.82 mm) while retaining biologically interpretable parameters such as asymptotic size and intrinsic growth rate [12, 13]. Although it was outperformed by ensemble and leading machine learning methods, it still produced an 80.5% RMSE reduction relative to its baseline form, showing that biological structure combined with environmental information provides substantial predictive power.

The contrast between mechanistic and ML approaches highlights a fundamental trade-off between interpretability and flexibility. ML models captured complex nonlinear relationships but rely on feature importance rather than explicit biological parameters [19, 20]. Mechanistic models provide direct biological meaning but exhibit greater variability in predictive performance. Hybrid ensemble approaches reconcile these strengths by combining mechanistic grounding with ML adaptability.

Limitations

The absence of consistent individual identifiers across capture events required treating each capture as independent, limiting the use of repeated measures or hierarchical growth models [21]. Environmental covariates were estimated at the reach level and may not fully reflect microhabitat conditions experienced by individual fish, particularly for temperature and discharge. The modeling framework was developed within the specific hydrological and ecological context of the Lower Colorado River, and its performance and optimal configuration will depend on species specific life history traits, data availability, and environmental drivers when applied to other systems. Growth observations were also unevenly distributed across size classes and seasons, which may influence model learning in sparsely sampled life stages.

Because individual trout could not be reliably linked across capture events, some degree of pseudo replication is possible if the same fish appears multiple times. This would primarily affect uncertainty estimates by inflating the effective sample size and

narrowing confidence or posterior intervals. However, the goal of this study was comparative prediction and out of sample performance rather than inference on population parameters. In this setting, repeated individuals do not systematically bias point predictions, and model rankings based on held out test error remain valid. Machine learning methods and Bayesian hierarchical models are also robust to moderate violations of independence because they learn conditional response surfaces rather than relying on asymptotic sampling theory. Thus, while future studies with individual tracking could refine uncertainty estimates, the observed performance advantages of the hybrid ensemble are unlikely to be artifacts of pseudo replication. To further assess the potential impact of pseudo replication, we conducted sensitivity checks comparing model performance across alternative data partitions and resampling schemes, which confirmed that the relative ranking and dominance of the hybrid ensemble were stable.

The stacked and Bayesian ensemble methods are also computationally more expensive than single model approaches, which may limit their direct use in real time management settings without adequate computational infrastructure or pre trained deployment pipelines.

Overall, these results support a framework in which ensemble methods integrating mechanistic and ML models provide the most reliable ecological forecasts. The 70–80% RMSE reductions relative to baseline models correspond to biologically meaningful improvements of 45–70 mm (20–32% of mean fish length). Even the smaller gains from ensemble integration, while modest in absolute RMSE, yield greater consistency across conditions, which is critical for management applications.

Conclusion

This study developed a unified, probabilistic framework for forecasting rainbow trout growth by integrating mechanistic biological models with modern machine–learning and ensemble methods. Across all metrics and posterior dominance comparisons, stacked ensembles achieved the most accurate and reliable predictions, while covariate-augmented mechanistic models (VBGM and Gompertz) remained competitive and retained essential biological interpretability. Top-performing models reduced RMSE by 70–80% relative to baseline approaches without covariates, demonstrating that environmental context fundamentally alters growth predictability rather than providing incremental gains.

These results show that trout growth in the Lower Colorado River is governed jointly by intrinsic physiological constraints and nonlinear environmental drivers, and that hybrid ensemble frameworks are uniquely capable of capturing both. By combining machine-learning accuracy, biological structure, and Bayesian uncertainty quantification, this approach provides managers with a robust and transferable tool for forecasting invasive species dynamics and evaluating policy interventions under uncertainty.

Glossary

- **Artificial Neural Network (ANN)**: A computational model composed of layered interconnected units (neurons) that learns nonlinear mappings between inputs and outputs through weighted combinations and nonlinear activation functions.
- **Bayesian Inference**: A statistical framework in which probability distributions are used to represent uncertainty about parameters and are updated via Bayes' theorem as new data are observed.
- **Bayesian Model Averaging (BMA)**: A Bayesian framework for prediction and inference in which multiple candidate models are combined by weighting each model's contribution according to its posterior probability given the data. Rather than selecting a single best model, BMA accounts for model uncertainty by averaging predictions and parameter estimates across the full model set, yielding posterior predictive distributions that integrate both parameter and structural uncertainty.
- **Cross-Validation (CV)**: A resampling procedure in which data are repeatedly partitioned into training and validation subsets to estimate out-of-sample predictive performance.
- Δt: Time elapsed between initial and final measurements.
- **Feature Importance**: A measure of the relative contribution of each input variable to a model's predictions, reflecting how strongly changes in a feature influence the predicted outcome. Feature importance is used to interpret model behavior, identify key drivers of prediction, and assess which biological, environmental, or observational factors most strongly shape the modeled response.
- **Gradient Boosting (GB)**: A technique that builds models sequentially to minimize the residual errors of previous models.
- **Hyperparameter Tuning**: The process of selecting algorithm control parameters (such as learning rate, tree depth, or regularization strength) that are not learned from the data but strongly influence model performance.
- K (Growth Coefficient): Rate parameter in the VBGM indicating how quickly the organism approaches its maximum size.
- L_1: Measured length of a fish at its first observation or tagging.
- L_2: Predicted length of a fish at a later time point.
- L_∞: Theoretical maximum length an organism can reach.
- **Machine Learning (ML)**: A class of computational methods that learn functional relationships from data to make predictions or decisions without explicit rule-based programming.
- **Mean Absolute Error (MAE)**: The average absolute difference between predicted and observed values.
- **No-U-Turn Sampler (NUTS)**: An adaptive Hamiltonian Monte Carlo algorithm that automatically tunes step size and path length to efficiently sample from complex posterior distributions without manual parameter tuning.

- **Overfitting**: A modeling error where the model captures noise instead of the underlying pattern.
- **Predictive Congruence**: The degree to which different models produce statistically and practically similar predictions for the same inputs.
- **Predictive Modeling**: The use of statistical or machine learning methods to forecast outcomes from data.
- **Random Forest (RF)**: An ensemble learning method using multiple decision trees to improve prediction accuracy.
- **Root Mean Square Error (RMSE)**: A standard measure of prediction error magnitude in models.
- **Stacked Ensemble**: An ensemble learning framework in which multiple base models are trained on the same data and their outputs are combined by a secondary model (the meta-learner) that learns how to optimally weight and integrate their predictions. By exploiting complementary strengths across different modeling approaches, stacking often achieves higher predictive accuracy and greater robustness than any individual constituent model.
- **Support Vector Regression (SVR)**: A machine learning method for predicting continuous outcomes using margin-based optimization.
- **Tag-Recapture Study**: A method to estimate growth or survival by tagging individuals and recapturing them later.
- **von Bertalanffy Growth Model (VBGM)**: A biological growth function describing the length of an organism as a function of age, commonly used in fisheries science.

Declarations

Use of Generative Artificial Intelligence

During the preparation of this work the authorsused ChatGPT4o for editing content. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

During the preparation of this work the authors used Claude.ai for code debugging. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Competing Interests

Authors have no competing interests.

Data Availability

The underlying USGS datasets are publicly available as cited. All code and results are available here: https://github.com/dustoff06/Trout.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author Contributions

P.L.: Conceptualization, Methodology, Software, Writing Original Draft. L.F.: Methodology, Software, Supervision, Writing Original Draft.

Supporting Information S1: Mathematical Specification

Data and notation

Let $i = 1, \ldots, N$ index tagged rainbow trout. For each individual we observe initial fork length $L_1^{(i)}$, recapture fork length $L_2^{(i)}$, elapsed time at large $t^{(i)}$, a vector of continuous covariates $\mathbf{X}_c^{(i)} \in \mathbb{R}^{p_c}$ (e.g., temperature, discharge), and a vector of categorical indicator variables $\mathbf{X}_d^{(i)} \in \{0,1\}^{p_d}$ for seasonal effects. The full dataset is

$$\mathcal{D} = \{L_1^{(i)}, L_2^{(i)}, t^{(i)}, \mathbf{X}_c^{(i)}, \mathbf{X}_d^{(i)}\}_{i=1}^N.$$

The modeling objective is to predict $L_2^{(i)}$ given $(L_1^{(i)}, t^{(i)}, \mathbf{X}_c^{(i)}, \mathbf{X}_d^{(i)})$.

Von Bertalanffy growth model with covariates

The von Bertalanffy growth model (VBGM) in Fabens formulation models the expected length at recapture for individual i as

$$\hat{L}_2^{(i)} = L_1^{(i)} + (L_\infty - L_1^{(i)})(1 - \exp(-k^{(i)} t^{(i)})),$$

equivalently written as

$$\hat{L}_2^{(i)} = L_\infty - (L_\infty - L_1^{(i)}) \exp(-k^{(i)} t^{(i)}),$$

where L_∞ is the asymptotic maximum fork length and $k^{(i)}$ is the individual-specific growth coefficient.

Covariates enter through a log-linear model for $k^{(i)}$:

$$k^{(i)} = \exp\Big(\alpha_k + \mathbf{X}_c^{(i)\top} \boldsymbol{\beta}_k + \mathbf{X}_d^{(i)\top} \boldsymbol{\beta}_{\text{dummies}}\Big).$$

Observed lengths are modeled as

$$L_2^{(i)} \mid \theta \sim \mathcal{N}(\hat{L}_2^{(i)}, \sigma_{\text{obs}}^2),$$

with parameter vector

$$\theta = \{L_\infty, \sigma_{\text{obs}}, \alpha_k, \boldsymbol{\beta}_k, \boldsymbol{\beta}_{\text{dummies}}\}.$$

Priors are specified as

$$\begin{aligned} L_\infty &\sim \mathcal{N}(500, \sigma_{L_\infty}^2), \quad \sigma_{L_\infty} \sim \mathcal{HN}(2) \\ \sigma_{\text{obs}} &\sim \mathcal{HN}(2) \\ \alpha_k &\sim \mathcal{N}(0,1) \\ \boldsymbol{\beta}_k &\sim \mathcal{N}(0,1) \\ \boldsymbol{\beta}_{\text{dummies}} &\sim \mathcal{N}(0, 0.2) \end{aligned}$$

The posterior is

$$p(\theta \mid \mathcal{D}) \propto \prod_{i=1}^{N} \mathcal{N}(L_2^{(i)} \mid \hat{L}_2^{(i)}, \sigma_{\text{obs}}^2)\, p(L_\infty)\, p(\sigma_{\text{obs}})\, p(\alpha_k)\, p(\boldsymbol{\beta}_k)\, p(\boldsymbol{\beta}_{\text{dummies}}).$$

Gompertz growth model with covariates

The Gompertz model provides an alternative mechanistic growth formulation with expected length at recapture

$$\hat{L}_2^{(i)} = L_\infty \exp\Big(-\exp(-k^{(i)} t^{(i)})\Big),$$

where L_∞ and $k^{(i)}$ are defined identically to the VBGM. Covariates enter through the same log-linear specification:

$$k^{(i)} = \exp\Big(\alpha_k + \mathbf{X}_c^{(i)\top}\boldsymbol{\beta}_k + \mathbf{X}_d^{(i)\top}\boldsymbol{\beta}_{\text{dummies}}\Big).$$

The likelihood, priors, and posterior follow the same structure as the VBGM.

Bayesian linear model

The Bayesian linear model assumes a normal likelihood with a linear predictor:

$$\begin{aligned} L_2^{(i)} &\sim \mathcal{N}(\mu^{(i)}, \sigma^2) \\ \mu^{(i)} &= \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij}^{(i)} \end{aligned}$$

where $x_{ij}^{(i)}$ is the value of the jth predictor for observation i, β_0 is the intercept, and $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_p\}$ are regression coefficients.

Priors are specified as

$$\beta_j \sim \mathcal{N}(0, 5) \text{ for } j = 0, 1, \ldots, p, \quad \log(\sigma) \sim \mathcal{N}(0, 1).$$

Machine learning predictors

Each machine-learning model m (Random Forest, XGBoost, LightGBM, SVR, ANN) defines a prediction function

$$\hat{L}_2^{(m,i)} = f_m(L_1^{(i)}, t^{(i)}, \mathbf{X}_c^{(i)}, \mathbf{X}_d^{(i)}),$$

where f_m is learned from the training data using cross-validation and hyperparameter tuning as described in the Methods.

For example, SVR with kernel function $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \phi(\mathbf{x}_j)$ solves

$$\min_{\mathbf{w}, b, \xi_i, \xi_i^*} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*)$$

subject to

$$\begin{aligned} L_2^{(i)} - \mathbf{w}^\top \phi(\mathbf{x}_i) - b &\le \varepsilon + \xi_i, \\ \mathbf{w}^\top \phi(\mathbf{x}_i) + b - L_2^{(i)} &\le \varepsilon + \xi_i^*, \\ \xi_i, \xi_i^* &\ge 0, \end{aligned}$$

with feature map $\phi(\cdot)$, regularization parameter C, and ε-insensitive loss margin.

Bayesian model averaging

For a set of candidate models $\mathcal{M}$, Bayesian Model Averaging (BMA) forms predictive distributions

$$p(\tilde{L}_2 \mid \mathcal{D}) = \sum_{m \in \mathcal{M}} p(\tilde{L}_2 \mid m, \mathcal{D})\, p(m \mid \mathcal{D}),$$

where

$$p(m \mid \mathcal{D}) \propto p(\mathcal{D} \mid m) p(m)$$

are posterior model probabilities. For models without closed-form likelihoods (e.g., tree-based ensembles), we approximate $p(\mathcal{D} \mid m)$ using cross-validation performance metrics or employ stacking weights as a pseudo-BMA approximation.

Posterior predictive errors

Let $\mathcal{T}$ denote the held-out test set of size N_T. For posterior draw s from model m, define

$$\mathrm{RMSE}_m^{(s)} = \sqrt{\frac{1}{N_T} \sum_{i \in \mathcal{T}} \big(L_2^{(i)} - \hat{L}_2^{(m,s,i)}\big)^2}, \qquad \mathrm{MAE}_m^{(s)} = \frac{1}{N_T} \sum_{i \in \mathcal{T}} \big|L_2^{(i)} - \hat{L}_2^{(m,s,i)}\big|.$$

For deterministic machine learning models (Random Forest, XGBoost, LightGBM, SVR, ANN), $S = 1$ and these reduce to point estimates. For Bayesian models (VBGM, Gompertz, Bayesian Linear), S equals the number of posterior samples, allowing full uncertainty quantification.

Stochastic dominance probabilities

The posterior probability that model a outperforms model b is

$$P(a \prec b \mid \mathcal{D}) = \frac{1}{S} \sum_{s=1}^{S} \mathbb{I}\Big(\mathrm{RMSE}_a^{(s)} < \mathrm{RMSE}_b^{(s)}\Big),$$

where S is the number of posterior samples (or $S = 1$ for deterministic models) and $\mathbb{I}(\cdot)$ is the indicator function. These probabilities define the entries in Table 3.

References

1. Diagne C, Leroy B, Vaissière AC, Gozlan RE, Roiz D, Jarić I, et al. High and rising economic costs of biological invasions worldwide. Nature. 2021;592(7855):571-6. doi:10.1038/s41586-021-03405-6.

2. Korman J. Early life history dynamics of rainbow trout in a large regulated river. University of British Columbia; 2009. doi:10.14288/1.0066949.

3. Yard MD, Korman J, Walters CJ, Kennedy TA. Seasonal and spatial patterns of growth of Rainbow Trout in the Colorado River in Grand Canyon, Arizona. Canadian Journal of Fisheries and Aquatic Sciences. 2016;73(1):125-39. doi:10.1139/cjfas-2015-0102.

4. Congressional Sportsmen's Foundation. Arizona secures U.S. Fish and Wildlife Service funding to support sport fish stocking program; 2021. Available from: `https://congressionalsportsmen.org/news/arizona-secures-us-fish-and-wildlife-service-funding-to-support-sport-fish-stocking-program/`.

5. Dutta H. Growth in fishes. Gerontology. 1994;40(2-4):97-112. doi:10.1159/000213581.

6. Sumpter JP. Control of growth of rainbow trout (Oncorhynchus mykiss). Aquaculture. 1991;100(1-3):299-320. doi:10.1016/0044-8486(92)90386-Y.

7. Korman J, Yard MD, Yackulic CB. Factors controlling the abundance of rainbow trout in the Colorado River in Grand Canyon in a reach utilized by endangered humpback chub. Canadian Journal of Fisheries and Aquatic Sciences. 2015;73(1):105-24. doi:10.1139/cjfas-2015-0101.

8. Korman J, Kaplinski M, Melis TS. Effects of fluctuating flows and a controlled flood on incubation success and early survival rates and growth of age-0 rainbow trout in a large regulated river. Transactions of the American Fisheries Society. 2011;140(2):487-505. doi:10.1080/00028487.2011.572015.

9. Harvey BC, Nakamoto RJ, White JL. Reduced streamflow lowers growth of rainbow trout in a small stream. Transactions of the American Fisheries Society. 2006;135(4):998-1005. doi:10.1577/T05-233.1.

10. Fontagné S, Silva N, Bazin D, Ramos A, Aguirre P, Surget A, et al. Effects of dietary phosphorus and calcium level on growth and skeletal development in rainbow trout (Oncorhynchus mykiss) fry. Elsevier. 2009;297(1-4):141-50. doi:10.1016/j.aquaculture.2009.09.022.

11. Rodehutscord M. Response of rainbow trout (Oncorhynchus mykiss) growing from 50 to 200 g to supplements of dibasic sodium phosphate in a semipurified diet. Elsevier. 1996;126(1):324-31. doi:10.1093/jn/126.1.324.

12. Fabens AJ. Properties and fitting of the von Bertalanffy growth curve. Growth. 1965. PMID: 5865688.

13. von Bertalanffy L. A quantitative theory of organic growth (inquiries on growth laws. II). JSTOR. 1938;10(2):181-213. doi:10.2307/41447359.

14. Gompertz B. XXIV. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. In a letter to Francis Baily, Esq. FRS &c. Philosophical transactions of the Royal Society of London. 1825;(115):513-83. doi:10.1098/rstl.1825.0026.

15. Ho L, Goethals P. Machine learning applications in river research: Trends, opportunities and challenges. Methods in ecology and evolution. 2022;13(11):2603-21. doi:10.1111/2041-210X.13992.

16. Flores A, Wiff R, Donovan CR, Gálvez P. Applying machine learning to predict reproductive condition in fish. Ecological Informatics. 2024;80:102481. doi:10.1016/j.ecoinf.2024.102481.

17. Isguzar S, Turkoglu M, Atessahin T, Durrani O. FishAgePredictioNet: A multi-stage fish age prediction framework based on segmentation, deep convolution network, and Gaussian process regression with otolith images. Fisheries Research. 2024;271:106916. doi:10.1016/j.fishres.2023.106916.

18. Li H, Chen Y, Li W, Wang Q, Duan Y, Chen T. An adaptive method for fish growth prediction with empirical knowledge extraction. Biosystems Engineering. 2021;212:336-46. doi:10.1016/j.biosystemseng.2021.11.012.

19. Pineda-Metz SE, Merk V, Pogoda B. A machine learning model and biometric transformations to facilitate European oyster monitoring. Aquatic Conservation: Marine and Freshwater Ecosystems. 2023;33(7):708-20. doi:10.1002/aqc.3912.

20. Bhai KVS, Ahmad SM, Gupta KYN, Jyothi C, Yunus M. Forecasting the Environment Factors for Optimal Plant Growth using XGBoost. In: 2024 2nd International Conference on Sustainable Computing and Smart Systems (ICSCSS). IEEE; 2024. p. 71-8. doi:10.1109/ICSCSS60660.2024.10625077.

21. Korman J, Yard MD. Effects of environmental covariates and density on the catchability of fish populations and interpretation of catch per unit effort trends. Fisheries Research. 2017;189:18-34. doi:10.1016/j.fishres.2017.01.005.

22. Python Software Foundation. Python Language Reference; 2025. Version 3.13. Available from: `https://docs.python.org/3/`.

23. Harris CR, Millman KJ, Van Der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585(7825):357-62. doi:10.1038/s41586-020-2649-2.

24. McKinney W, et al. Data structures for statistical computing in Python. In: Proceedings of the 9th Python in Science Conference; 2010. p. 51-6. doi:10.25080/Majora-92bf1922-00a.

25. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine learning in Python. the Journal of Machine Learning Research. 2011;12:2825-30. doi:10.5555/1953048.2078195.

26. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785-94. doi:10.1145/2939672.2939785.

27. Ke G, Meng Q, Finley T, Wang T, Chen W, Ma W, et al. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In: Guyon I, von Luxburg U, Bengio S, Wallach H, Fergus R, Vishwanathan S, et al., editors. Advances in Neural Information Processing Systems 30 (NeurIPS 2017); 2017. p. 3146-54. Available from: `https://proceedings.neurips.cc/paper_files/paper/2017/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf`.

28. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. arXiv preprint arXiv:191201703. 2019. doi:10.5555/3454287.3455008.

29. Brown G, Wyatt JL, Tiňo P. Managing Diversity in Regression Ensembles. Journal of Machine Learning Research. 2005;6:1621-50. Available from: `https://jmlr.org/papers/v6/brown05a.html`.

30. Phan D, Pradhan N, Jankowiak M. Composable effects for flexible and accelerated probabilistic programming in NumPyro; 2019. ArXiv:1912.11554. Available from: `http://arxiv.org/abs/1912.11554`.

31. Hoffman MD, Gelman A, et al. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. Journal of Machine Learning Research. 2014;15(1):1593-623. Available from: `https://www.jmlr.org/papers/volume15/hoffman14a/hoffman14a.pdf`.

32. Drucker H, Burges CJC, Kaufman L, Smola A, Vapnik V. Support Vector Regression Machines. In: Advances in Neural Information Processing Systems. vol. 9. MIT Press; 1996. p. 155-61. doi:10.5555/2998981.2999003.

33. Breiman L. Random forests. Machine Learning. 2001;45(1):5-32. doi:10.1023/A:1010933404324.

34. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2016. p. 785-94. doi:10.1145/2939672.2939785.

35. McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. The Bulletin of Mathematical Biophysics. 1943;5(4):115-33. doi:10.1007/BF02478259.