Title: Predicting Migration Survival in Swainson's Thrushes Using Machine Learning

Supplementary Materials: Supp Mat

1. Introduction

Ecological speciation is the process through which divergent adaptation contributes to the evolution of reproductive isolation between two or more species (Rundle & Nosil, 2005; Schluter, 2009). This form of reproductive isolation can evolve via the accumulation of extrinsic post-zygotic barriers, which are expressed when hybrid fitness is reduced in nature (Coyne & Orr, 2004). A key requirement for demonstrating extrinsic postzygotic isolation is connecting ecologically relevant traits that are under divergent selection to reductions in hybrid fitness. This can be challenging. For starters, estimating fitness in natural populations is notoriously difficult (Orr, 2009). Even if fitness can be quantified, it can be difficult to connect specific traits to reductions in fitness because incipient species have often diverged in many traits, generating correlations between them (Gillespie et al., 2018; Losos, 1990; Schluter & McPhail, 1992). Hybrid zones, regions where incipient species come into contact and hybridize, can help to address this (Hewitt, 1988; Thompson et al., 2024). Hybridization and subsequent recombination can generate a broader range of trait values as well as highly variable combinations of traits, which offers an opportunity to observe how these trait combinations perform in nature (Marques, Meier, et al., 2019).

Although hybridization offers an opportunity to observe and connect patterns of trait and fitness variation, the nature of relationships between traits and fitness can still make them challenging to characterize. Fitness landscapes can be complex, with multiple local maxima (Martin & Wainwright, 2013; Pfaender et al., 2016). Generally, the parental species' phenotypes

are expected to have high fitness, but this may not be true in hybrid zone environments, and there may also be novel trait combinations or intermediate phenotypes that perform well (Lamichhaney et al., 2017; Marques, Lucek, et al., 2019; Meier et al., 2017). Furthermore, traits may have interactive effects on fitness, so that the effect of one trait on fitness depends on others (Thompson et al., 2024). Therefore, it can be difficult to disentangle which traits are important for fitness variation in hybrid zones, particularly if interactive effects are non-linear or if several traits are involved.

In the present study, we used a hybrid zone formed by two subspecies of Swainson's thrushes, combined with a modelling approach based in machine learning, to overcome challenges associated with identifying traits that contribute to ecological speciation. The two Swainson's thrush subspecies form a migratory divide in the Coast Mountains of western North America; the inland subspecies migrates via eastern North America to Colombia while the coastal subspecies migrates to Central America along the Pacific coast (Delmore et al., 2012; K. C. Ruegg & Smith, 2002). As in other examples of migratory divides, seasonal migration is hypothesized to act as an extrinsic, post-zygotic barrier in the Swainson's thrush (Alvarado et al., 2014; Scordato et al., 2020; Veen et al., 2007). Consistent with this hypothesis, hybrid routes are ecologically inferior to parental routes (Justen et al. 2020) and both hybrids and coastal backcrosses survive seasonal migration at lower rates than the parental species (Blain et al., 2024). Therefore, we estimated survival on migration, quantified by individual tracking with automated radio telemetry (Delmore & Easton, 2019; Taylor et al., 2017), as a proxy for fitness. The subspecies have diverged in a suite of traits that are associated with long-distance migration, ranging from behaviours, such as migratory timing, to morphological traits, including wing shape (Delmore et al., 2012; K. Ruegg, 2008). Hybrids exhibit high variability and diverse

combinations of migratory traits (Delmore & Irwin, 2014). However, these traits do not show linear or quadratic relationships to survival in hybrids, suggesting that if there are links between traits and fitness, they are complex or interactive (Blain et al., 2024). Therefore, machine learning based approaches may be a valuable avenue through

Machine learning (ML) is a powerful tool for analyzing high-dimensional ecological data and revealing complex, non-linear and interactive relationships that traditional methods struggle to detect (Stupariu et al., 2022). In the context of Swainson's Thrush migration, ML is particularly effective for integrating diverse behavioral and morphological traits to understand how they interact to influence survival on migration. Our study uses ML to connect these traits to fitness, measured by survival on migration, addressing the challenges posed by their complex relationships. ML helps uncover patterns in trait interactions that might otherwise remain hidden, making it crucial for understanding the ecological dynamics of this migratory divide.

We applied two ML models, Random Forest and Neural Networks, to examine the link between trait variation and survival in Swainson's thrush hybrids. Random Forest, an ensemble method, creates multiple decision trees to predict outcomes, while Neural Networks, with interconnected layers of nodes, learn hierarchical data representations (Dongare et al., 2012; Haddouchi & Berrado, n.d.). Both models excel in handling large datasets and revealing intricate trait interactions that affect migration survival. By leveraging these techniques, we aim to identify key traits and interactions that shape the survival on migration in hybrid Swainson's thrushes, thereby linking ecologically-relevant traits and fitness to better understand postzygotic isolation in this system.

2. Methods

2.1. Data Collection

We captured Swainson's thrushes using mist nets in the Pemberton area of British Columbia, Canada. Juvenile birds were identified based on the presence of yellow tips on their wing coverts (n = 479). The captures occurred during the breeding seasons from August to September in the years 2019 through 2023, aligning with the period when juveniles are most active and preparing for migration. Each captured bird underwent morphological measurements, and we equipped them with nanotags—small radio transmitter devices that emit a unique signal – to measure migratory behaviours and quantify survival. These signals are detected by an array of radio antennas deployed across North and Central America, enabling us to track the birds' movements throughout the year. We also collected small blood samples from all individuals for genetic analysis. Blood samples were taken back to the lab for DNA extraction and sequencing to identify the nucleotide sequences and alleles present in each bird's genome.

2.2. Phenotypic Features

Upon capture, multiple morphological measurements were taken. Kipp's distance, a measure of wing pointedness, was measured as the length between the longest primary feather and the first secondary feather, while distal wing length was measured as the distance from p10 to the longest feather (mm). Additional measurements included tarsus length (mm), weight (g), tail length (mm), feather lengths (p7, p8, p9, p10; mm) and wing cord length (mm), which represent critical aspects of structural morphology. Body condition was computed as the ratio of weight to tarsus length (g/mm), providing an index of the bird's overall health. Three behavioral variables were analyzed: release day, migratory timing, and orientation. Release day was the day of the year that a bird was caught and tagged while migrating through the hybrid zone. Fall

detect day 1 was defined as the day of the year a bird was first detected by a Motus radio station, within 30 km of its release site, during fall migration. Fall bearing 1 was calculated as the orientation of the bird's movement during fall migration, measured from the release site to the first Motus station where the bird was detected, within a 30 km radius. Both variables were derived from data recorded by the Motus network of radio telemetry stations.

2.3. Genotypic Features

Blood samples were collected for genetic analysis, yielding ancestry and heterozygosity data. Ancestry was estimated based on the proportion of inland versus coastal alleles, ranging from 0 for entirely coastal birds to 1 for entirely inland birds. Heterozygosity, the proportion of loci that were heterozygous, reflected genetic diversity, with values ranging from 0 for fully homozygous birds to 1 for F1 hybrids.

2.4. Normalization

For the neural network model, we normalized the predictor variables to a range of 0 to 1 to ensure consistent scaling across all inputs, which improves convergence during model training (Cabello-Solorzano et al., 2023). This was achieved using a custom scaling function in R:

$$\text{rangeNorm}(x) = \frac{x - \min(x, \text{na.rm} = \text{TRUE})}{\max(x, \text{na.rm} = \text{TRUE}) - \min(x, \text{na.rm} = \text{TRUE})}$$

Here, x represents the variable to be normalized, and na.rm = TRUE ensures that missing values are ignored when computing the minimum and maximum. This method scales each variable to fall between 0 and 1 without altering its statistical properties, making the data suitable for neural network architectures, which are sensitive to the scale of input features.

For the Random Forest model, we used the raw, unnormalized predictor values because Random Forests are scale-invariant and do not require feature normalization (Cabello-Solorzano et al., 2023). This approach preserves the original distributions of the predictor variables, allowing the Random Forest to naturally handle varying scales and ranges.

2.5. Preprocessing

Several features were highly correlated with each other, introducing redundancy into our analyses (Figure S4a). To address this, we removed one feature from pairs with correlation coefficients of 0.7 or higher, such as p7, p8, and p9, which were feather measurements that were all highly correlated with wing cord, and weight, which was correlated with body condition (Rickert et al., 2023). Additionally, some features had significant amounts of missing data. We removed features with more than 25% missing data (e.g., carpal length and fat score, excluding fall detect day 1 and fall bearing 1 due to its significance in analysis; Figure S4b) and used K-Nearest Neighbors (KNN) imputation to handle missing values in the remaining features (Halder et al., 2024; Shadbahr et al., 2023). KNN imputation was chosen because it estimates missing values based on the average of the values of the “k” nearest neighbors, preserving the characteristics of similar observations and making it well-suited for ecological data. These "nearest neighbors" are determined by the distance between samples in a multi-dimensional space defined by the dataset’s features (here, trait values). We selected k=5 to balance smoothing out noise while capturing relevant patterns (Figure S4c).

The target variable, survival, presented a significant class imbalance, with 78.7% of samples representing birds that did not survive migration and only 21.3% representing survivors.

This imbalance could lead the model to predict non-survival for all samples and still achieve misleadingly high accuracy (Figure S5a). To address this, we applied the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic samples for the minority class (Sahlaoui et al., 2024). Just as with KNN imputation, we used k = 5, so that the generation of synthetic samples occurs by interpolating between a randomly selected minority class sample and its five nearest neighbors (Figure S5). SMOTE was particularly suitable for our dataset due to its ability to preserve natural variability in ecological traits. By generating synthetic samples close to actual data points, SMOTE maintained the relationships among bird traits without oversampling identical values, thereby enhancing the model's capacity to learn patterns from both survival and non-survival outcomes.

2.6. Random Forest Model

We used a Random Forest model to predict survival on migration in Swainson's thrushes. Random Forest is an ensemble learning method that constructs multiple decision trees during training and combines their predictions to improve accuracy and reduce overfitting (Haddouchi & Berrado, n.d.). Each tree is built using a bootstrap sample of the data, and at each node, a random subset of predictors is considered for splitting. This randomness ensures that the model is robust to noise and less prone to overfitting.

The final prediction is determined by aggregating the outputs of all trees through majority voting for classification tasks. For our study, we configured the Random Forest in Python using the RandomForestClassifier from the scikit-learn library (Testas, 2023). Key hyperparameters include n_estimators (the number of trees in the forest), max_depth (the maximum depth of each tree), min_samples_split (the minimum number of samples required to split an internal node),

min_samples_leaf (the minimum number of samples required to be at a leaf node), and max_features (the number of features to consider when looking for the best split). These hyperparameters are tuned to optimize the model's performance. To optimize the model's performance, we used GridSearchCV from scikit-learn to perform an exhaustive search over specified hyperparameter values. We evaluated model fits with accuracy scores, confusion matrices, and feature importance rankings to identify the most influential predictors of survival (see below).

2.7. Neural Network Model

Neural networks are a class of machine learning models inspired by the structure of the human brain, consisting of interconnected layers of nodes (neurons) (Dongare et al., 2012). Each connection between neurons has an associated weight that is adjusted during training to minimize prediction error. Our model was a multi-layer perceptron (MLP), a feedforward neural network composed of an input layer, one or more hidden layers, and an output layer. The input layer received the predictor variables, the hidden layers applied non-linear activation functions to capture complex relationships between predictors, and the output layer produced the final survival predictions.

The model was trained using MLPClassifier from the scikit-learn library in python (Li et al., 2022). Key hyperparameters include hidden_layer_sizes (the number of neurons in each hidden layer), activation (the activation function for the hidden layers, e.g., 'relu', 'tanh'), solver (the optimization algorithm, e.g., 'adam', 'sgd'), alpha (the regularization term to prevent overfitting), and learning_rate (the learning rate schedule for weight updates). These were tuned to optimize model performance using the same method as earlier. Additionally, the model

evaluation includes the same metrics as previously described, but feature importance is replaced with permutation importance.

2.8. Training

We split the data randomly into two datasets, with 80% for training and 20% for testing (Joseph & Vakayil, 2022). Each model was trained on 80% of the dataset, which included both trait values and survival outcomes. This training process enabled the model to learn how to distinguish between traits associated with survival and those linked to non-survival. To evaluate performance, the model made predictions on the unseen 20% test set. We compared predicted and actual outcomes to assess accuracy.

2.9. Using Confusion Matrix for Evaluation

A confusion matrix is a performance evaluation tool that summarizes the predictions of a classification model by comparing the predicted labels with the actual labels (Düntsch & Gediga, 2019). It consists of four key metrics: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). True positives and true negatives represent correctly classified instances of the positive and negative classes, respectively, while false positives indicate instances incorrectly classified as positive, and false negatives reflect instances incorrectly classified as negative. These metrics allow for the calculation of performance measures such as accuracy, precision, recall (sensitivity), specificity, and the F1-score.

2.10. Using Feature/Permutation Importance and Pairwise Importance for Evaluation

The feature importance plot for the Random Forest model highlights the relative contribution of each feature to the model's predictive performance (Menze et al., 2009). It is created by measuring the decrease in model accuracy (or increase in impurity) when the feature is randomly shuffled, disrupting its relationship with the target variable. Features that cause a significant drop in model performance when shuffled are deemed more important. The importance score is computed as the mean decrease in impurity (MDI) across all trees in the forest. By examining the feature importance plot, we can identify which features drive the model's predictions, providing insights into model interpretability.

Feature importance in neural networks can be assessed using permutation importance, which measures the change in model performance when a feature's values are randomly shuffled (Altmann et al., 2010). The permutation importance plot for a neural network evaluates the impact of each feature on the model's performance by measuring the change in the model's accuracy when the values of a feature are randomly shuffled. This disruption breaks the relationship between the feature and the target variable, and the resulting decrease in performance reflects the importance of that feature in making accurate predictions.

Although both methods measure the impact of each feature, the results may differ due to the underlying differences in model architectures. However, both plots provide valuable insights into which features most influence model predictions and can be used to compare the importance of different features across models.

Pairwise importance quantifies the extent to which interactions between two features contribute to the predictive performance of a model. Unlike individual feature importance, which measures the isolated contribution of a single feature, pairwise importance captures the additional predictive value gained when the features are considered together (Loh & Zhou,

2021). For example, high pairwise importance may indicate that the relationship between the two features is synergistic. Interpretation involves identifying feature pairs with the highest interaction importance and assessing whether these interactions align with domain-specific hypotheses or indicate patterns requiring further exploration. In practice, these values are especially relevant for understanding model behavior in contexts where feature dependencies are biologically or scientifically meaningful.

2.11. Using SHAP for Evaluation

SHAP (Shapley Additive Explanations) plots provide a powerful tool for interpreting machine learning model predictions by quantifying the contribution of each feature to the model's output, offering insights into specific features' contributions to survival predictions (Zhao et al., 2020). 1D SHAP dependence plots present the relationship between trait values and their influence on predictions. In these plots, each dot in the SHAP plots represents a sample, with its position along the X-axis showing the feature's actual value and the Y-axis showing the SHAP value (its contribution to the prediction). SHAP values near zero indicate negligible influence on predictions, while values further from zero indicate stronger impacts. A clear, monotonic trend in the plot (e.g., a positive or negative slope) suggests that the feature has a consistent and predictable influence on the model's predictions. Conversely, a more scattered or non-linear distribution of points indicates that the feature's effect on the prediction may be complex or dependent on other factors.

3. Results

The results of the Random Forest model are summarized in Figure 1. The bar plot (Figure 1A) displays feature importances derived from the model using the mean decrease in impurity (Gini index). A red dashed line represents the threshold of 1/number of features, above which features are considered to have relatively high importance. Key features exceeding this threshold include tail length, ancestry, fall bearing 1, kipps, fall detect day 1, body condition, and heterozygosity. The interaction importance table (Figure 1B) highlights the top feature pairs exceeding the threshold of 1/number of possible combinations, indicating how feature interactions jointly influence the model's predictions. Figure 1C presents 1D SHAP dependence plots for the most important features identified in the bar plot. These plots illustrate the relationship between each feature and the model's predicted outcome, showing how changes in feature values impact individual predictions.

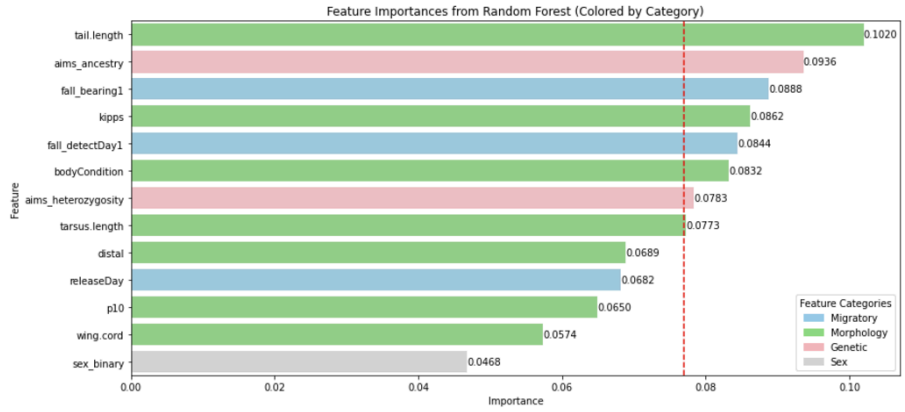
Model performance is evaluated using the confusion matrix (Figure 1D), which shows 94 true negatives, 16 false positives, 17 false negatives, and 99 true positives, yielding an accuracy of 85%.

For the neural network model, feature importance was determined using permutation scores (Figure 2A). Features with relatively high importance include kipps, fall bearing 1, wing cord, ancestry, and heterozygosity.

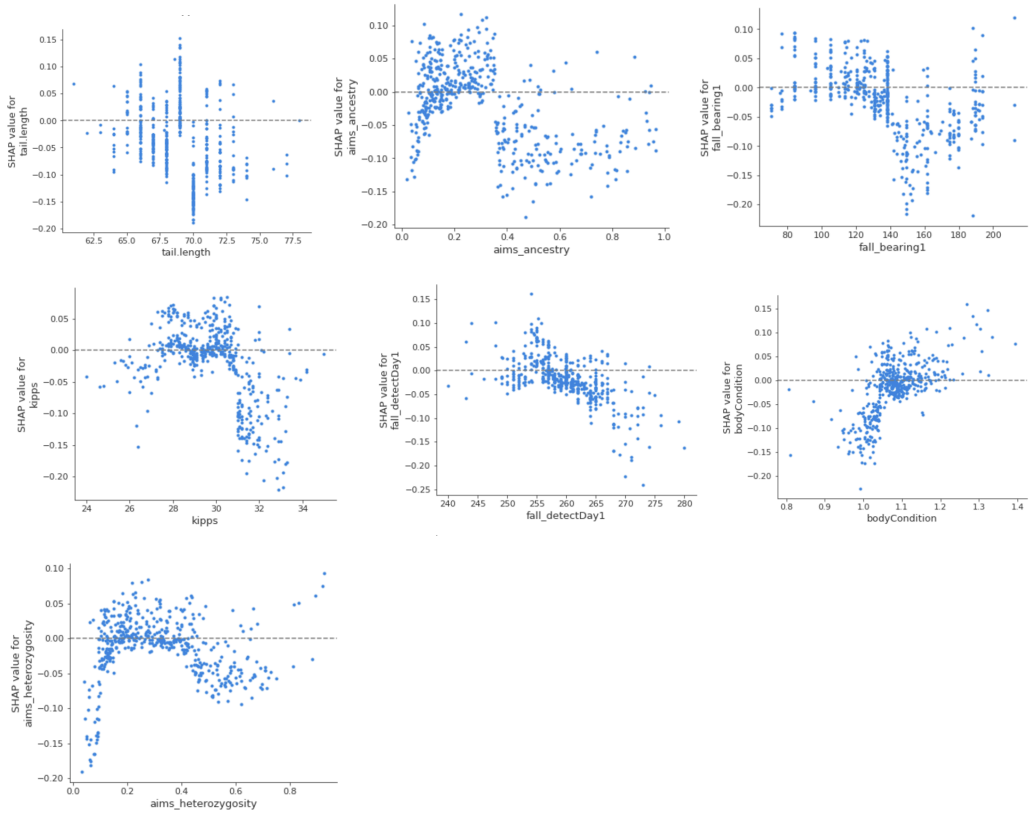
The confusion matrix for the neural network model (Figure 2B) reports 76 true negatives, 34 false positives, 23 false negatives, and 93 true positives, resulting in an accuracy of 75%.

Figure 1: Random Forest Model Evaluation

(A) Feature Importance Bar Plot



(B) SHAP 1D Plots for Top Features



(C) Top Pairwise Importance Scores

feature_1	feature_2	importance
tail.length	sex_binary	0.0192
fall_detectDay1	sex_binary	0.0181
fall_bearing1	tail.length	0.0167
fall_bearing1	distal	0.0166
fall_detectDay1	fall_bearing1	0.0162
kipps	sex_binary	0.0161
fall_detectDay1	wing.cord	0.0160
aims_heterozygosity	aims_ancestry	0.0158
tarsus.length	sex_binary	0.0156
fall_bearing1	kipps	0.0156
fall_bearing1	wing.cord	0.0155
fall_bearing1	releaseDay	0.0155
fall_bearing1	p10	0.0149
fall_bearing1	tarsus.length	0.0144
bodyCondition	sex_binary	0.0142
tarsus.length	tail.length	0.0139
tarsus.length	aims_ancestry	0.0137
releaseDay	aims_ancestry	0.0137
fall_bearing1	bodyCondition	0.0136
sex_binary	aims_ancestry	0.0135
wing.cord	sex_binary	0.0135
sex_binary	releaseDay	0.0135
fall_detectDay1	tail.length	0.0135
kipps	releaseDay	0.0135
distal	sex_binary	0.0134
fall_detectDay1	aims_ancestry	0.0134
distal	p10	0.0132
tail.length	aims_ancestry	0.0132
p10	aims_ancestry	0.0131
kipps	aims_ancestry	0.0130
fall_detectDay1	kipps	0.0130
sex_binary	aims_heterozygosity	0.0129
bodyCondition	releaseDay	0.0128

(D) Confusion Matrix

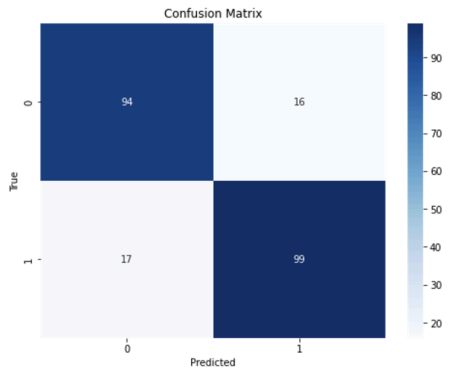
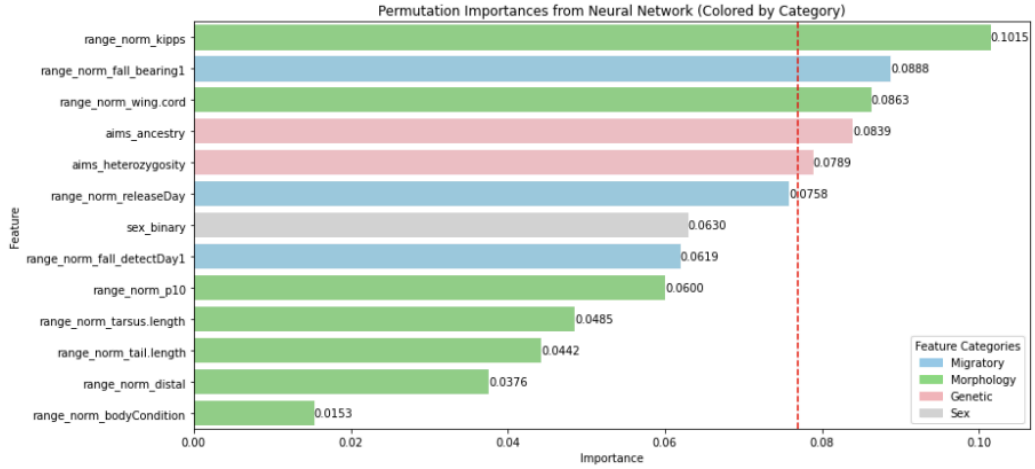
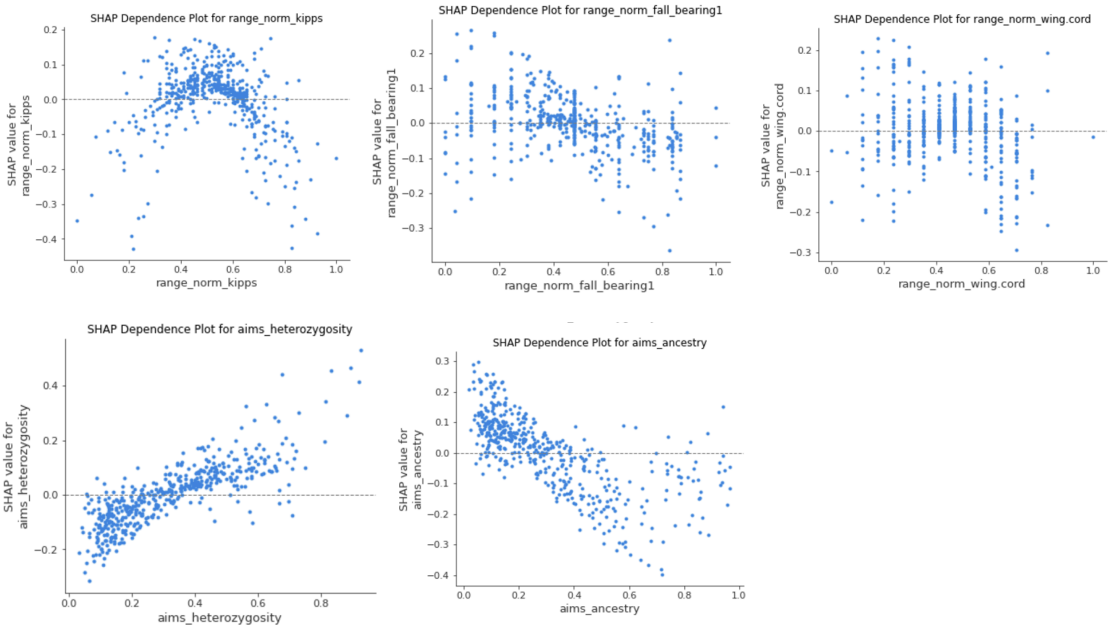


Figure 2: Neural Network Model Evaluation

(A) Permutation Importance Bar Plot



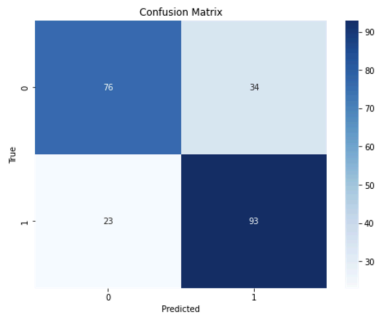
(B) SHAP 1D Plots for Top Features



(C) Top Pairwise Importance Scores

feature_1	feature_2	importance
range_norm_fall_detectDay1	sex_binary	0.0293
range_norm_fall_bearing1	sex_binary	0.0271
range_norm_p10	sex_binary	0.0236
sex_binary	range_norm_releaseDay	0.0229
range_norm_tarsus.length	aims_ancestry	0.0216
range_norm_tarsus.length	sex_binary	0.0204
range_norm_releaseDay	aims_ancestry	0.0201
sex_binary	aims_ancestry	0.0192
range_norm_wing.cord	sex_binary	0.0191
range_norm_distal	sex_binary	0.0180
range_norm_tail.length	sex_binary	0.0179
range_norm_fall_bearing1	range_norm_releaseDay	0.0172
range_norm_bodyCondition	sex_binary	0.0172
range_norm_kipps	sex_binary	0.0165
range_norm_fall_detectDay1	aims_ancestry	0.0163
range_norm_fall_bearing1	range_norm_tarsus.length	0.0159
range_norm_fall_detectDay1	range_norm_fall_bearing1	0.0156
range_norm_fall_bearing1	range_norm_distal	0.0155
range_norm_fall_bearing1	range_norm_wing.cord	0.0150
range_norm_fall_bearing1	aims_ancestry	0.0150
range_norm_wing.cord	range_norm_releaseDay	0.0148
range_norm_distal	range_norm_p10	0.0147
range_norm_fall_detectDay1	aims_heterozygosity	0.0145
range_norm_kipps	range_norm_releaseDay	0.0144
sex_binary	aims_heterozygosity	0.0143
range_norm_distal	range_norm_wing.cord	0.0143
range_norm_fall_bearing1	range_norm_p10	0.0142
range_norm_kipps	aims_ancestry	0.0137
range_norm_tail.length	aims_ancestry	0.0136
range_norm_p10	range_norm_wing.cord	0.0135
range_norm_tail.length	range_norm_releaseDay	0.0133
range_norm_distal	range_norm_releaseDay	0.0133
range_norm_p10	range_norm_tail.length	0.0132
range_norm_fall_detectDay1	range_norm_releaseDay	0.0131

(D) Confusion Matrix



References

Altmann, A., Toloşi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, *26*(10), 1340–1347. https://doi.org/10.1093/bioinformatics/btq134

Alvarado, A. H., Fuller, T. L., & Smith, T. B. (2014). Integrative tracking methods elucidate the evolutionary dynamics of a migratory divide. *Ecology and Evolution*, *4*(17), 3456–3469. https://doi.org/10.1002/ece3.1205

Blain, S. A., Justen, H. C., Easton, W. E., & Delmore, K. E. (2024). Reduced hybrid survival in a migratory divide between songbirds. *Ecology Letters*, *27*, e14420.

Cabello-Solorzano, K., Ortigosa De Araujo, I., Peña, M., Correia, L., & J. Tallón-Ballesteros, A. (2023). The Impact of Data Normalization on the Accuracy of Machine Learning Algorithms: A Comparative Analysis. In P. García Bringas, H. Pérez García, F. J. Martínez De Pisón, F. Martínez Álvarez, A. Troncoso Lora, Á. Herrero, J. L. Calvo Rolle, H. Quintián, & E. Corchado (Eds.), *18th International Conference on Soft Computing Models in Industrial and Environmental Applications (SOCO 2023)* (Vol. 750, pp. 344–353). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-42536-3_33

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Sinauer Associates.

Delmore, K. E., & Easton, W. (2019). *BC Interior Thrushes (Project280)*. Motus Wildlife Tracking System, Birds Canada. https://motus.org/

Delmore, K. E., Fox, J. W., & Irwin, D. E. (2012). Dramatic intraspecific differences in migratory routes, stopover sites and wintering areas, revealed using light-level geolocators. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1747), 4582–4589. https://doi.org/10.1098/rspb.2012.1229

Delmore, K. E., & Irwin, D. E. (2014). Hybrid songbirds employ intermediate routes in a migratory divide. *Ecology Letters*, *17*(10), 1211–1218. https://doi.org/10.1111/ele.12326

Dongare, A. D., Kharde, R. R., & Kachare, A. D. (2012). *Introduction to Artificial Neural Network*. *2*(1).

Düntsch, I., & Gediga, G. (2019). Confusion matrices and rough set data analysis. *Journal of Physics: Conference Series*, *1229*(1), 012055. https://doi.org/10.1088/1742-6596/1229/1/012055

Gillespie, R. G., Benjamin, S. P., Brewer, M. S., Rivera, M. A. J., & Roderick, G. K. (2018). Repeated diversification of ecomorphs in Hawaiian stick spiders. *Current Biology*, *28*(6), 941-947.e3. https://doi.org/10.1016/j.cub.2018.01.083

Haddouchi, M., & Berrado, A. (n.d.). *A survey and taxonomy of methods interpreting random forest models*.

Halder, R. K., Uddin, M. N., Uddin, Md. A., Aryal, S., & Khraisat, A. (2024). Enhancing K-nearest neighbor algorithm: A comprehensive review and performance analysis of modifications. *Journal of Big Data*, *11*(1), 113. https://doi.org/10.1186/s40537-024-00973-y

Hewitt, G. M. (1988). Hybrid zones-natural laboratories for evolutionary studies. *Trends in Ecology & Evolution*, *3*(7), 158–167. https://doi.org/10.1016/0169-5347(88)90033-X

Joseph, V. R., & Vakayil, A. (2022). SPlit: An Optimal Method for Data Splitting. *Technometrics*, *64*(2), 166–176. https://doi.org/10.1080/00401706.2021.1921037

Lamichhaney, S., Han, F., Webster, M. T., Andersson, L., Grant, R. B., & Grant, P. R. (2017). Rapid hybrid speciation in Darwin's finches. *Science*, *4593*.

Li, W., Chen, H., Guo, J., Zhang, Z., & Wang, Y. (2022). Brain-inspired Multilayer Perceptron with Spiking Neurons. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 773–783. https://doi.org/10.1109/CVPR52688.2022.00086

Loh, W.-Y., & Zhou, P. (2021). Variable Importance Scores. *Journal of Data Science*, 569–592. https://doi.org/10.6339/21-JDS1023

Losos, J. B. (1990). Ecomorphology, performance capability, and scaling of West Indian Anolis lizards: An evolutionary analysis. *Ecological Monographs*, *60*(3), 369–388.

Marques, D. A., Lucek, K., Sousa, V. C., Excoffier, L., & Seehausen, O. (2019). Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12182-w

Marques, D. A., Meier, J. I., & Seehausen, O. (2019). A combinatorial view on speciation and adaptive radiation. *Trends in Ecology and Evolution*, *34*(6), 531–544. https://doi.org/10.1016/j.tree.2019.02.008

Martin, C. H., & Wainwright, P. C. (2013). Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science*, *339*(6116), 208–211. https://doi.org/10.1126/science.1227710

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, *8*(14363), 1–11. https://doi.org/10.1038/ncomms14363

Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., & Hamprecht, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC Bioinformatics*, *10*(1), 213. https://doi.org/10.1186/1471-2105-10-213

Orr, H. A. (2009). Fitness and its role in evolutionary genetics. *Nature Reviews Genetics*, *10*(8), 531–539. https://doi.org/10.1038/nrg2603

Pfaender, J., Hadiaty, R. K., Schliewen, U. K., & Herder, F. (2016). Rugged adaptive landscapes shape a complex, sympatric radiation. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1822), 11–13. https://doi.org/10.1098/rspb.2015.2342

Rickert, C. A., Henkel, M., & Lieleg, O. (2023). An efficiency-driven, correlation-based feature elimination strategy for small datasets. *APL Machine Learning*, *1*(1), 016105. https://doi.org/10.1063/5.0118207

Ruegg, K. (2008). Genetic, morphological, and ecological characterization of a hybrid zone that spans a migratory divide. *Evolution*, *62*(2), 452–466.

https://doi.org/10.1111/j.1558-5646.2007.00263.x

Ruegg, K. C., & Smith, T. B. (2002). Not as the crow flies: A historical explanation for circuitous migration in Swainson's thrush (*Catharus ustulatus*). *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *269*(1498), 1375–1381. https://doi.org/10.1098/rspb.2002.2032

Rundle, H. D., & Nosil, P. (2005). Ecological speciation. *Ecology Letters*, *8*, 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Sahlaoui, H., Alaoui, E. A. A., Agoujil, S., & Nayyar, A. (2024). An empirical assessment of smote variants techniques and interpretation methods in improving the accuracy and the interpretability of student performance models. *Education and Information Technologies*, *29*(5), 5447–5483. https://doi.org/10.1007/s10639-023-12007-w

Schluter, D. (2009). Evidence for ecological speciation and its alternative. *Science*, *323*(5915), 737–741. https://doi.org/10.1126/science.1160006

Schluter, D., & McPhail, J. D. (1992). Ecological character displacement and speciation in sticklebacks. *The American Naturalist*, *140*(1), 85–108.

Scordato, E. S. C., Smith, C. C. R., Semenov, G. A., Liu, Y., Wilkins, M. R., Liang, W., Rubtsov, A., Sundev, G., Koyama, K., Turbek, S. P., Wunder, M. B., Stricker, C. A., & Safran, R. J. (2020). Migratory divides coincide with reproductive barriers across replicated avian hybrid zones above the Tibetan Plateau. *Ecology Letters*, *23*, 231–241. https://doi.org/10.1111/ele.13420

Shadbahr, T., Roberts, M., Stanczuk, J., Gilbey, J., Teare, P., Dittmer, S., Thorpe, M., Torné, R. V., Sala, E., Lió, P., Patel, M., Preller, J., AIX-COVNET Collaboration, Selby, I., Breger, A., Weir-McCall, J. R., Gkrania-Klotsas, E., Korhonen, A., Jefferson, E., … Schönlieb, C.-B. (2023). The impact of imputation quality on machine learning classifiers for datasets with missing values. *Communications Medicine*, *3*(1), 139. https://doi.org/10.1038/s43856-023-00356-z

Stupariu, M.-S., Cushman, S. A., Pleşoianu, A.-I., Pătru-Stupariu, I., & Fürst, C. (2022). Machine learning in landscape ecological analysis: A review of recent approaches. *Landscape Ecology*, *37*(5), 1227–1250. https://doi.org/10.1007/s10980-021-01366-9

Taylor, P. D., Crewe, T. L., Mackenzie, S. A., Lepage, D., Aubry, Y., Crysler, Z., Finney, G., Francis, C. M., Guglielmo, C. G., Hamilton, D. J., Holberton, R. L., Loring, P. H., Mitchell, G. W., Norris, D. R., Paquet, J., Ronconi, R. A., Smetzer, J. R., Smith, P. A., Welch, L. J., & Woodworth, B. K. (2017). The Motus Wildlife Tracking System: A collaborative research network. *Avian Conservation & Ecology*, *12*(1), 8.

Testas, A. (2023). *Distributed Machine Learning with PySpark: Migrating Effortlessly from Pandas and Scikit-Learn*. Apress. https://doi.org/10.1007/978-1-4842-9751-3

Thompson, K. A., Brandvain, Y., Coughlan, J. M., Delmore, K. E., Justen, H., Linnen, C. R., Ortiz-Barrientos, D., & Rushworth, C. A. (2024). The ecology of hybrid incompatibilities. *Cold Spring Harbor Perspectives in Biology*, a041440. https://doi.org/10.1101/cshperspect.a041440

Veen, T., Svedin, N., Forsman, J. T., Hjernquist, M. B., Hjernquist, K. A. T., Träff, J., Klaassen, M., Veen, T., Svedin, N., Forsman, J. T., Hjernquist, M. B., Qvarnstr, A., Hjernquist, K. A. T., Tr, J., & Klaassen, M. (2007). Does migration of hybrids contribute to post-zygotic isolation in flycatchers? *Proceedings of the Royal Society B: Biological Sciences*, *274*(1610), 707–712. https://doi.org/10.1098/rspb.2006.0058

Zhao, W., Joshi, T., Nair, V. N., & Sudjianto, A. (2020). *SHAP values for Explaining CNN-based Text Classification Models*.