Machine learning identifies ecological selectivity patterns across the end-Permian mass extinction

Authors: Foster, William J., Ayzel, Georgy, Münchmeyer, Jannes, Rettelbach, Tabea, Kitzmann, Niklas H., et al.

PALEOBIOLOGY
A PUBLICATION OF THE

Paleontological
S O C I E T Y

Article

Machine learning identifies ecological selectivity patterns across the end-Permian mass extinction

William J. Foster* [iD], Georgy Ayzel [iD], Jannes Münchmeyer [iD], Tabea Rettelbach [iD], Niklas H. Kitzmann [iD], Terry T. Isson, Maria Mutti, and Martin Aberhan

Abstract.—The end-Permian mass extinction occurred alongside a large swath of environmental changes that are often invoked as extinction mechanisms, even when a direct link is lacking. One way to elucidate the cause(s) of a mass extinction is to investigate extinction selectivity, as it can reveal critical information on organismic traits as key determinants of extinction and survival. Here we show that machine learning algorithms, specifically gradient boosted decision trees, can be used to identify determinants of extinction as well as to predict extinction risk. To understand which factors led to the end-Permian mass extinction during an extreme global warming event, we quantified the ecological selectivity of marine extinctions in the well-studied South China region. We find that extinction selectivity varies between different groups of organisms and that a synergy of multiple environmental stressors best explains the overall end-Permian extinction selectivity pattern. Extinction risk was greater for genera that had a low species richness, narrow bathymetric ranges limited to deep-water habitats, a stationary mode of life, a siliceous skeleton, or, less critically, calcitic skeletons. These selective losses directly link the extinctions to the environmental effects of rapid injections of carbon dioxide into the ocean–atmosphere system, specifically the combined effects of expanded oxygen minimum zones, rapid warming, and potentially ocean acidification.

William J. Foster. Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany. University College Dublin, Belfield, Dublin 4, Ireland. *Present address: Universität Hamburg, Institute for Geology, Hamburg, Germany. E-mail: w.j.foster@gmx.co.uk

Georgy Ayzel. Universität Potsdam, Institute for Geosciences, Potsdam-Golm, Germany. E-mail: ayzel@uni-potsdam.de

Jannes Münchmeyer. GFZ German Research Centre for Geosciences, Potsdam, Germany. Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, Germany. E-mail: munchmej@gfz-potsdam.de

Tabea Rettelbach. Alfred Wegener Institute Helmholtz Centre for Polar and Marine Research, Potsdam, Germany. Universität Potsdam, Institute for Geosciences, Potsdam-Golm, Germany. Humboldt-Universität zu Berlin, Department of Computer Science, Berlin, Germany. E-mail: tabea.rettelbach@awi.de

Niklas H. Kitzmann. Potsdam Institute for Climate Impact Research (PIK)—Member of the Leibniz Association, Potsdam, Germany. Universität Potsdam, Institute of Physics and Astronomy, Potsdam-Golm, Germany. E-mail: kitzmann@pik-potsdam.de

Terry T. Isson. University of Waikato, School of Science, Tauranga, New Zealand. E-mail: terry.isson@waikato.ac.nz

Maria Mutti. Universität Potsdam, Institute for Geosciences, Potsdam-Golm, Germany. E-mail: mmutti@geo.uni-potsdam.de

Martin Aberhan. Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany. E-mail: Martin.Aberhan@mfn.berlin

Introduction

The end-Permian mass extinction event occurred approximately 252 Ma during one of the most extreme climate warming events of the Phanerozoic (Burgess and Bowring 2015). The timing of this extinction event overlaps with the emplacement of the Siberian Traps large igneous province into carbon-rich sedimentary rocks and the associated injection of large volumes of greenhouse gases into the atmosphere (Svensen et al. 2009). Based on our understanding of the impact of rapid and

high-magnitude carbon dioxide emissions, as well as both sedimentological and geochemical proxies from the Permian/Triassic boundary, it has been hypothesized that the causes of marine extinctions stem from ocean acidification (e.g., Payne et al. 2007), thermal stress (e.g., Joachimski et al. 2012), and deoxygenation (e.g., Wignall and Twitchett 1996). Even though there is evidence that these environmental changes rapidly developed over the same time interval as the extinction event, it still must be shown that the environmental changes directly contributed to the extinctions in the oceans.

One way to explore the cause(s) of a mass extinction event is to investigate the ecological extinction selectivity. Several biological and ecological traits appear to have been selected against during the end-Permian mass extinction. Specifically, selectivity patterns of extinction have been recognized among an organism's physiology (Knoll et al. 2007; Clapham and Payne 2011; Payne et al. 2016), mineralogy (Clapham and Payne 2011), motility (Clapham 2017), and tiering (Clapham and Payne 2011). These selective patterns have led to suggestions that decreasing pH, hypercapnia, and global ocean warming played a major role in the extinctions of marine organisms (Knoll et al. 2007; Clapham and Payne 2011; Payne et al. 2016; Clapham 2017). Furthermore, a biogeographic selectivity pattern has been related to the metabolic rate of marine organisms, suggesting that thermal stress in the tropics and deoxygenation toward the poles best explain the pattern of extinction (Penn et al. 2019). It is, however, unlikely that the end-Permian extinction selected equally strongly against all these different ecological attributes, and some of these traits appear selected against because many of the traits are shared among the taxa that went extinct. In addition, extinction selectivity is highly complex, and these previous studies use statistical approaches that do not reveal key information about interactions among these traits.

An alternative approach to investigate extinction selectivity is to utilize machine learning algorithms, which can model interdependencies between features. Machine learning is one of the major subfields of artificial intelligence, classically defined as computational algorithms that have the ability to learn from data as to improve performance or make accurate predictions (Mohri et al. 2018). Machine learning techniques can be subdivided into supervised learning (labeled input and output variables and an algorithm that produces an inferred function to make predictions), unsupervised learning (algorithms infer a function to describe a hidden structure from unlabeled data), or semi-supervised (algorithm is trained upon a combination of labeled and unlabeled data) learning techniques. Statistical methods and machine learning algorithms are not mutually exclusive, but "statistics draws population inferences from a sample, and machine learning finds generalizable predictive patterns" (Bzdok et al. 2018: p. 1). Machine learning algorithms also typically compromise interpretability for predictive power, which makes statistical methods a better choice for revealing relationships within the data. State-of-the-art advances in machine learning, however, mean that various methods have now been proposed to help users interpret the predictions of complex models (Lundberg and Lee 2017), which also improves their utility in extinction selectivity studies.

In this study, we apply decision tree algorithms, some of the simplest, but powerful, supervised machine learning algorithms, which can solve both regression and classification problems (Breiman et al. 1984). Decision trees are a set of nodes (conditions) and branches derived from the data, with the terminal nodes (leaf nodes) representing classification outputs/ decisions similar to a flowchart diagram. Decision tree algorithms aim to find which conditions return the greatest information gain based on evaluating certain metrics and then pick a combination of conditions that gives the highest/lowest value of a specific metric (Breiman et al. 1984). Decision tree algorithms offer immense potential to contribute to problem solving in geosciences (Karapatne et al. 2017) and can also be used to predict survival during mass mortality events. A key example is the Titanic dataset (Kaggle 2022), used to train data scientists in predicting survivorship based upon the metadata of passenger information (e.g., gender, wealth, age). Despite the potential of decision tree algorithms to predict extinction risk and its many applications across

biology (e.g., Boyer 2010; Hanna and Cardillo 2014), there have been few studies that have utilized this technique for paleobiology (Finnegan et al. 2012, 2015; Novack-Gottshall 2016; Smith et al. 2018; Tietje and Rödel 2018). Besides generating a model to predict extinction risk, decision tree algorithms can also be used to rank ecological factors that are most important when making the predictive model and thus can reveal extinction selectivity patterns. To date, decision tree algorithms have only been utilized for exploring extinction selectivity during one mass extinction (i.e., the end-Ordovician mass extinction; Finnegan et al. 2012); this study used a random forest algorithm in which a large number of decision trees are generated and combined using averages at the end of the process. A different approach is to use gradient boosted decision trees, where each decision tree is built upon the previous tree, rather than linearly combining trees at the end, which can result in a better performance (Natekin and Knoll 2013). To demonstrate the potential of decision tree algorithms in revealing extinction selectivity, we explore ecological selectivity during the end-Permian mass extinction using the categorized gradient boosting decision tree algorithm as implemented in the CatBoost software library (Dorogush et al. 2018). Here, through application of a decision tree algorithm, the relative importance of the determinants of extinction can be resolved, which will reveal the relative importance of explanatory traits for the end-Permian event. Furthermore, we also include the application of novel methods to increase the interpretability of the machine learning output.

To investigate extinction selectivity patterns across the end-Permian mass extinction, we examined the ecological traits that were selected against to infer extinction mechanisms. The intensely studied and species-rich rock record of marine successions from South China provides an excellent opportunity to compare selectivity patterns between preextinction and the extinction interval while minimizing geographic, rock-record, and sampling biases. It is particularly important that South China covers a broad region with numerous sections and facies representing time intervals before the mass extinction that also occur in other sections afterward, which allows for a consistency with both the type of marine environment and paleolatitude of the investigated Permian–Triassic successions. This limits the impact of a facies' control on fossil occurrence patterns and any sequence stratigraphic overprints on the extinction pattern. In addition, because 63% of pre-extinction Changhsingian and 41% of postextinction Griesbachian occurrences in the Paleobiology Database (paleobiodb.org) are derived from China, and because China is one of the few regions that records both a continuous and fossiliferous deposition across the Permian/Triassic boundary, this high volume of data for China biases our "global" understanding of the end-Permian mass extinction (Fig. 1). This region, therefore, offers the best perspective on extinction dynamics for this event.

Methods

We used a database of all genera of marine invertebrates, conodonts, and calcareous algae recorded from the Lopingian to the Middle Triassic in South China downloaded from the Paleobiology Database and supplemented with additional data from the literature (Supplementary Fig. S1). The database includes 25,683 occurrences of 1283 genera representing Brachiopoda, Foraminifera, Mollusca, Conodonta, Radiolaria, Arthropoda, Chlorophyta, Rhodophyta, Porifera, Cnidaria, Echinodermata, Bryozoa, and Annelida. Vetting of the data meant that undetermined genera and informal genera were excluded from the database. In addition, individual species were vetted to ensure that they were not represented within multiple genera in the database through taxonomic synonymy, and the most up-to-date generic identification of the species was followed. To estimate the number of species within a genus, the number of named species for each genus in each time bin was tabulated and $\log_e$ transformed. Occurrences without a species name or designated as indeterminate species (e.g., sp. or spp.) were assumed to represent a single species. Latest synonymies and reidentifications were used where possible, and for mollusks were updated following MolluscaBase (molluscabase.org). Given that
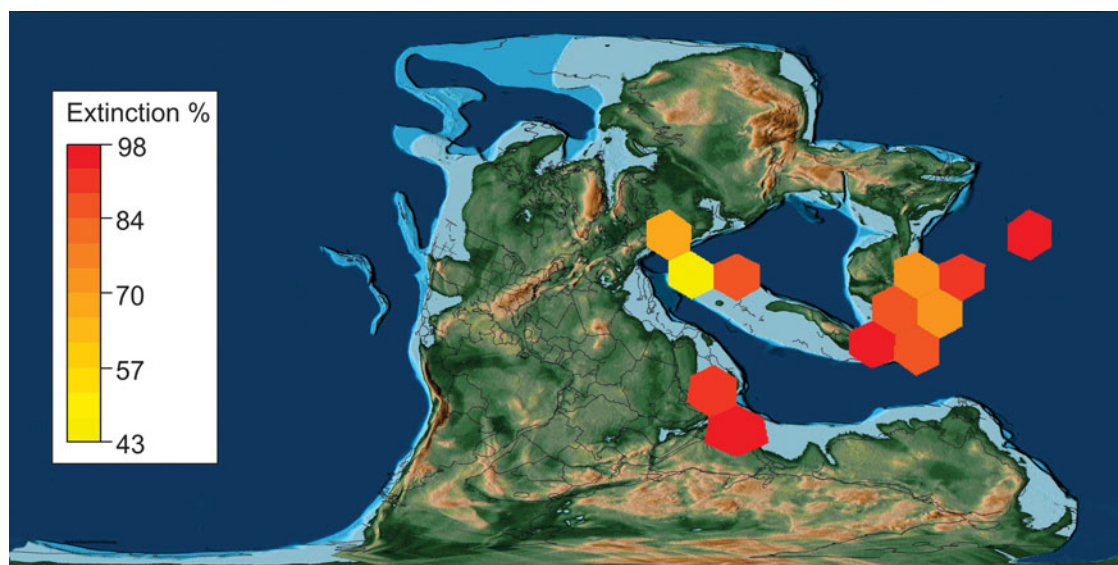
FIGURE 1. Proportional extinctions of marine genera during the end-Permian mass extinction event in different regions. Each hexagon cell represents an equal area, and only cells that include both pre- and postextinction data are included. The plot was generated using data from the Paleobiology Database (paleobiodb.org) and plotted using the icosa package (Kocsis et al. 2018) in R overlain on a base map from Scotese (2016).

metadata information (e.g., depositional environment, section name, and formation name) were sometimes inconsistent among data sources, even for the same fossiliferous bed in the same section, or because it was absent from the Paleobiology Database, the metadata were updated for each section using the most up-to-date literature for consistency. Occurrences of taxa from outside the South China region that extend the stratigraphic ranges of genera into the Permian or into the Triassic were not included to avoid introducing spatial variation and biases in the dataset. Where possible, using the original references, each occurrence was assigned to a conodont biozone (Supplementary Fig. S1). In some sections, in particular the Permian/Triassic Global Boundary Stratotype Section and Point in Meishan, single beds have been subsequently divided into subbeds that relate to different conodont biozones. Therefore, in older references that did not subdivide these beds, the occurrence is considered to have been present in all respective biozones. All analyses were carried out at the genus level because of increased likelihood of misidentifications at the species level, 29% of species occurrences in our database lack a formal species-level identification, and the

high proportion of singleton species. For all the analyses, genera with single occurrences were omitted. Extinction rates were calculated for those genera that could be assigned to conodont biozones, that is 12,889 occurrences of 634 genera. Extinction rates were calculated following Foote (2000): extinction rate = $-\log[N_{bt}/(N_{bL} + N_{bt})]$, where N_{bL} is the number of taxa that cross the bottom boundary of a time bin only and N_{bt} is the number of taxa that cross both boundaries.

To investigate extinction selectivity, we characterized each genus according to 10 ecological traits (Table 1) and two phylogenetic criteria (phylum and within-genus species richness) using the primary literature, extant relatives, and Paleobiology Database references to the ecological attribute during a taxon's adult life stage (see also Supplemental Material). These traits were chosen because they are expected to show selectivity signals that would test the role of the expected consequences of climate change and can be applied to all the investigated groups in this study. The geographic range of each taxon was not included in the analysis. This is because the Permian–Triassic fossil record is strongly biased to a few regions (see Fig. 1), which means that the reconstructed

TABLE 1. Ecological categories used in this study. Physiological buffering capacity after Knoll et al. (2007) and Payne et al. (2016). Tiering and motility after Bambach et al. (2007). Ornamentation after Aberhan et al. (2006). Reproduction mode after Bush et al. (2016). Descriptions for the remaining categories are given in the Supplementary Material. Physiology and shell mineralogy are ranked according to the expected sensitivity to ocean acidification following Knoll et al. (2007) and Ries (2011), respectively. *Bimineralic refers to calcareous skeletons that are a mixture of aragonite and calcite shell layers. [†]Taxa without a calcitic shell were classified as having a carbonate load of 0. Carbonate load is calculated as the body size multiplied by the $CaCO_3$ dry weight calculated for each class for modern representatives. [‡]Bathymetric range is calculated by the number of broad depositional settings from a taxon's minimum depth to their maximum depth.

Physiology	Tiering	Ornamentation
1. Heavy carbonate load and no buffering capacity	1. Pelagic	1. Smooth
2. Moderate carbonate load and moderate buffering capacity	2. Epifaunal	2. Fine
3. Little or no carbonate load	3. Semi-infaunal	3. Strong
	4. Infaunal	
Shell mineralogy		Bathymetric range[‡]
	Motility	1. One setting
1. Aragonite	1. Motile, fast	2. Two settings
2. High-/intermediate-Mg calcite	2. Motile, slow	3. Three settings
3. Low-Mg calcite	3. Facultatively motile	4. Four settings
4. Bimineralic*	4. Stationary, unattached	
5. Organophosphate	5. Stationary, attached	Phylum
6. Silica		1. Annelida
7. No shell	Reproductive mode	2. Arthropoda
	1. Broadcaster	3. Brachiopoda
Body size (mm)	2. Intermediate	4. Bryozoa
1. Micro (0.01–1)	3. Non-broadcaster	5. Chlorophyta
2. Meio (1.01–10)		6. Chordata
3. Macro (>10.01)		7. Cnidaria
	Respiratory protein	8. Echinodermata
	1. Other	9. Foraminifera
Carbonate load[†]	2. Hemoglobin	10. Mollusca
1. No carbonate (0)	3. Hemocyanin	11. Porifera
2. Low (0.01–1)	4. Hemerythrin	12. Problematica
3. Moderate (1.01–10)		13. Radiolaria
4. High (>10.01)		14. Rhodophyta

geographic range is overprinted by the limited number of fossiliferous Permian/Triassic boundary successions rather than representing an ecological signal. In addition, both Payne and Finnegan (2007) and Clapham and Payne (2011) demonstrated that geographic range is not a good predictor of extinction risk for the end-Permian mass extinction. Genera that exhibited ecological traits that differed between species, for example, differences in the strength of ornamentation, were split into different taxa to consider within-genus ecological variation. To analyze extinction selectivity, the data were binned into four time intervals due to the nature of the end-Permian mass extinction event (see "Results") and to avoid oversplitting the data: the Wuchiapingian, pre-extinction Changhsingian, the mass extinction interval (Changhsingian *Clarkina yini* to the Griesbachian *Hindeodus parvus* Zone; as defined by Wang et al. 2014), and the Griesbachian postextinction interval.

Categorical Gradient Boosting Decision Trees.— To identify and rank the key drivers of extinction risk, we applied the categorical gradient boosting decision trees (CatBoost) method. CatBoost is a powerful and high-performance machine learning technique with a better prediction performance in comparison with conventional linear models (Ayaru et al. 2015; Belyaeva et al. 2017). The core idea behind the CatBoost technique is to combine simple decision tree models with gradient boosting (Friedman 2001) in such a way that the next decision tree is combined with the previous decision tree and the weights of the classifiers are modified to minimize the error of the previous decision tree. In the present study, we used a CatBoost model to predict extinction risk as a binary classification problem, that is, whether a species went extinct ("1") or not ("0") based on the set of corresponding traits. To fit the CatBoost model, we used the open-source machine learning library for gradient boosting CatBoost

(Dorogush et al. 2018; Prokhorenkova et al. 2018). The feature importance of the CatBoost model was done using the PredictionValuesChange algorithm, wherein the feature importance values are normalized so that the sum of all features is equal to 100 (Dorogush et al. 2018).

Class imbalance (where the dataset is biased toward one of the classes) is an important issue to consider while dealing with classification problems. Here, the distribution of our data into two classes (extinct [1] and not extinct [0] samples) is relatively balanced: only for the Wuchiapingian are the data unbalanced, as the percentage of "extinct" samples is around 11%, while for the remaining time intervals this value fluctuates from 39% to 57%. To consider the class imbalance issue that could be present during the cross-validation, we use the stratified approach for data splitting, which ensures that analyzed folds preserve the percentage of samples for each class.

We trained an individual CatBoost model for each investigated time interval with default CatBoost hyperparameters, fivefold cross-validation (i.e., splits, where data are randomly split into a training and evaluation subset), and the log loss function as an optimization criterion. Machine learning algorithms can often be configured using a wide range of hyperparameters. Although the Catboost algorithm is known for its even performance across different hyperparameter choices, the effect of varying the more important hyperparameters should be evaluated. We tested the performance of a number of Catboost models created using different sets of hyperparameters (learning rate, L2 regularization, and tree depth). Significant differences between the performance on testing and training data can be discerned in runs with a high learning rate, pointing to overfitting. However, the performance on testing data is remarkably stable across a wide range of hyperparameters, justifying the choice of using the default parameters of the algorithm (Supplementary Fig. S2, Supplementary Table S1). This generally low impact of hyperparameters on the test performance can be attributed to several reasons: (1) gradient boosted trees are stable with respect to hyperparameter selection (Cawley and Talbot 2010); (2) the prediction task is complex due to the fact that not all taxa in a functional trait have the same status of extinct/not extinct, but the model can only assign one prediction; and (3) the significance of hyperparameter optimization in the study is also questionable as the datasets are small, where very small changes in the prediction already would have a considerable impact on the score. It is also worth noting that using multiple models and comparing their performance already provides a good indicator that the model does not underperform.

We estimated the obtained model performance with the area under the receiver operating characteristic (AUC) which is sensitive to type I and type II errors. An AUC is a measure of discrimination between two distinct groups of species we try to classify as extinctions or survivors; thus, it is closely related to the Wilcoxon signed-rank test, which determines whether two randomly dependent samples have the same distribution. An AUC was averaged over the splits we used for cross-validation. Additionally, we performed recursive feature elimination analysis (as proposed in Belyaeva et al. 2017), which confirmed the robustness of the proposed approach in ranking traits. As we expected correlations among our variables, we checked the predictors for multicollinearity before model building (Supplementary Fig. S3). A correlation plot of the different predictors shows that a few predictors are correlated, for example, body size and carbonate load during the extinction interval (Supplementary Fig. S3). Generally, correlation values are low, however, indicating that no undesirable levels of multicollinearity are present.

Machine learning algorithms are often described as a black box due to the lack of transparency associated with how algorithms make predictions. To demonstrate how the CatBoost algorithm generates its predictions and to increase interpretability, we applied Shapley additive explanations (SHAP) values to exemplify the output from the CatBoost algorithm (Lundberg and Lee 2017). First, we investigated the collective SHAP values for each variable, which highlight how much each predictor contributes to the target variable, and then the individual SHAP values for each genus, which show how the contribution of different features affects the prediction.

All the data and code regarding performed machine learning modeling and analysis are available on GitHub (https://github.com/PaleoML/permian-selectivity). Each part of our research workflow can be reproduced interactively using Binder.[1]

The results using the Catboost method were also compared with random forest and multivariate logistic regression algorithms. These comparisons showed that Catboost produced better AUC values by 0.08 and 0.03 than the random forest methods for the Wuchiapingian and Changhsingian, respectively, and an AUC value 0.07 better than the logistic regression for the extinction interval (Supplementary Fig. S4). In addition, the Catboost method (SHAP range: −0.8 to 0.7) produced a larger range of SHAP values than the random forest algorithm (SHAP range: −0.25 to 0.3), allowing clearer inferences about the importance of a given feature and within-feature selectivity patterns to be drawn (Supplementary Fig. S5).

Results

Nature of the End-Permian Mass Extinction.— Exploring the selectivity of the end-Permian mass extinction first requires the timing and the duration of the event to be determined. The mass extinction event has been described as a single event in the latest Permian (Jin et al. 2000), two separate events extending into the earliest Triassic (Song et al. 2013), or a single extended extinction interval spanning the Permian/Triassic boundary (Wang et al. 2014). Our analyses of extinction rates in South China show that those genera that can be correlated to conodont biozones have heightened extinction rates in the *Clarkina yini* Zone, in the *Hindeodus zhejiangensis–Hindeodus changxingensis* Zones, and throughout most of the Griesbachian (Fig. 2, Supplementary Fig. S6), but a low extinction rate in the *Clarkina meishanensis* Zone (Fig. 2). This low extinction rate in the *C. meishanensis* Zone also corresponds with a relatively low number of occurrences compared with surrounding zones, and the low extinction rate is likely a consequence of

insufficient sampling. Given that the highest extinction rates occurred in the extinction interval as defined by Wang et al. (2014), we follow those researchers and interpret the extinction event as representing an interval spanning the Permian/Triassic boundary, which also avoids oversplitting the data. This suggests that the extrinsic changes that occurred over this approximately 60 kyr interval (Burgess and Bowring 2015) caused the end-Permian mass extinction. Consequently, to investigate extinction selectivity, the data were aggregated into four time intervals: the Wuchiapingian, pre-extinction Changhsingian, the mass extinction interval (Changhsingian *C. yini* to the Griesbachian *Hindeodus parvus* Zone; Wang et al. 2014), and the Griesbachian postextinction interval.

Extinction Selectivity.—The AUC (Fig. 3) visualizes the CatBoost model performance. This curve plots the true-positive rate against the false-positive rate. An AUC of 0.5 indicates the performance of a random classification that has no utility (Fig. 3). It is unlikely that a decision tree algorithm will give an AUC value of 1, which represents a perfect classification; instead, an AUC > 0.7 is typically considered representative of a good model (Mohri et al. 2018). The AUCs for each time interval (Fig. 3) show an average of 0.75, 0.84, 0.81, and 0.72 for the Wuchiapingian, pre-extinction Changhsingian, mass extinction interval, and Griesbachian postextinction time interval, respectively. In particular, the curves for the pre-extinction Changhsingian and the mass extinction interval show that the CatBoost model is a good classification model for interpreting extinction selectivity. This means that the machine learning algorithm creates a good model for understanding extinction selectivity in the pre-extinction Changhsingian and extinction interval.

CatBoost reveals that the relative importance of the ecological variables associated with extinction varies between intervals (Fig. 4, Supplementary Table S2). The difference in the relative importance of factors between the pre-extinction Changhsingian and the extinction interval indicates that the end-Permian mass extinction was not simply triggered by an intensification of pre-extinction pressures, a characteristic shared with other mass extinction events (Finnegan et al.

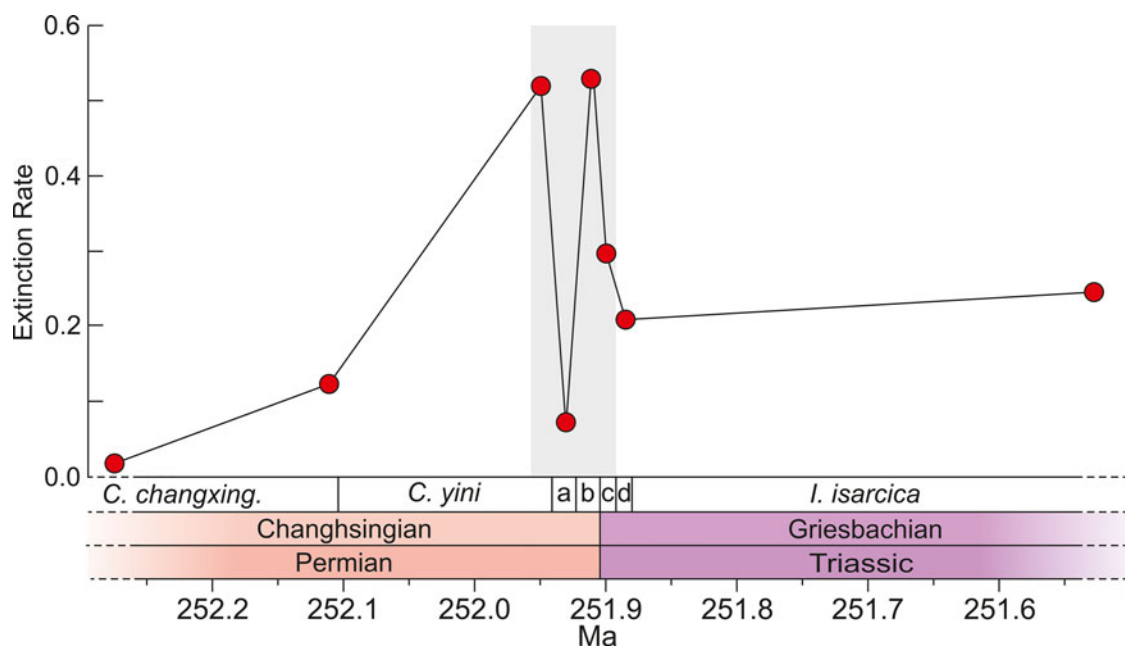[1]https://mybinder.org/v2/gh/PaleoML/permian-selectivity/HEAD.

FIGURE 2. High-resolution extinction rates of marine genera across the Permian/Triassic boundary in South China. The extinction interval is highlighted in gray after Wang et al. (2014). Radiometric ages after Burgess and Bowring (2015). Conodont zones after Yuan et al. (2014); a, *Clarkina meishanensis*; b, *Hindeodus zhejiangensis–Hindeodus changxingensis*; c, *Hindeodus parvus*; d, *Isarcicella staeschei, C. changxing, Clarkina changxingensis, I. isarcica = Isarcicella isarcica*. Extinction rates for the full studied interval are shown in Supplementary Fig. S6.

2012; Dunhill et al. 2018). The relative importance of different factors for the extinction interval highlights that bathymetric range, within-genus species richness, skeletal mineralogy, and physiology are the best predictors of extinction risk. These were also better predictors of extinction than phylum, showing that these ecological attributes are more significant in predicting extinction than phylogenetic membership. Bathymetric range and number of species were important predictors of extinction before the defined extinction interval (Fig. 4), suggesting that some of the main drivers of ecological selectivity before the extinction interval were still important during the extinction interval. There is, however, also a correlation between the number of occurrences and both the number of species and bathymetric range of a genus (Supplementary Fig. S3), which suggests that these variables are potentially also affected by sampling and taxonomic artifacts. In this case, the importance of species richness and bathymetric range as predictors of extinction risk might be slightly exaggerated in our results. Furthermore, when exploring the genera that the

algorithm misclassified, it became clear that misclassified genera had a low number of both species and occurrences. These genera are, therefore, interpreted to have a poor fossil record, which may explain their unpredictability.

The SHAP values improve the interpretability of feature importance, as they show how the features are used by the decision tree algorithms. The SHAP summary plot (Fig. 5) reveals the SHAP value for each genus for each feature. The values range between −0.8 and 0.7 and the further a value is from 0, the more influence it will have on the prediction, whereas values near 0 will have less of an impact on predictions. The SHAP summary plot (Fig. 5) therefore shows that phylum is a poor predictor of extinction, because most of the SHAP values are close to 0, and mineralogy is a good predictor of extinction, because clusters of samples with the same classification have large positive and negative values. The SHAP force plots (Fig. 6) show the sum of the SHAP values across all the features added to the base value (the average predicted probability) for selected genera, which is used by the algorithm to make its predictions.
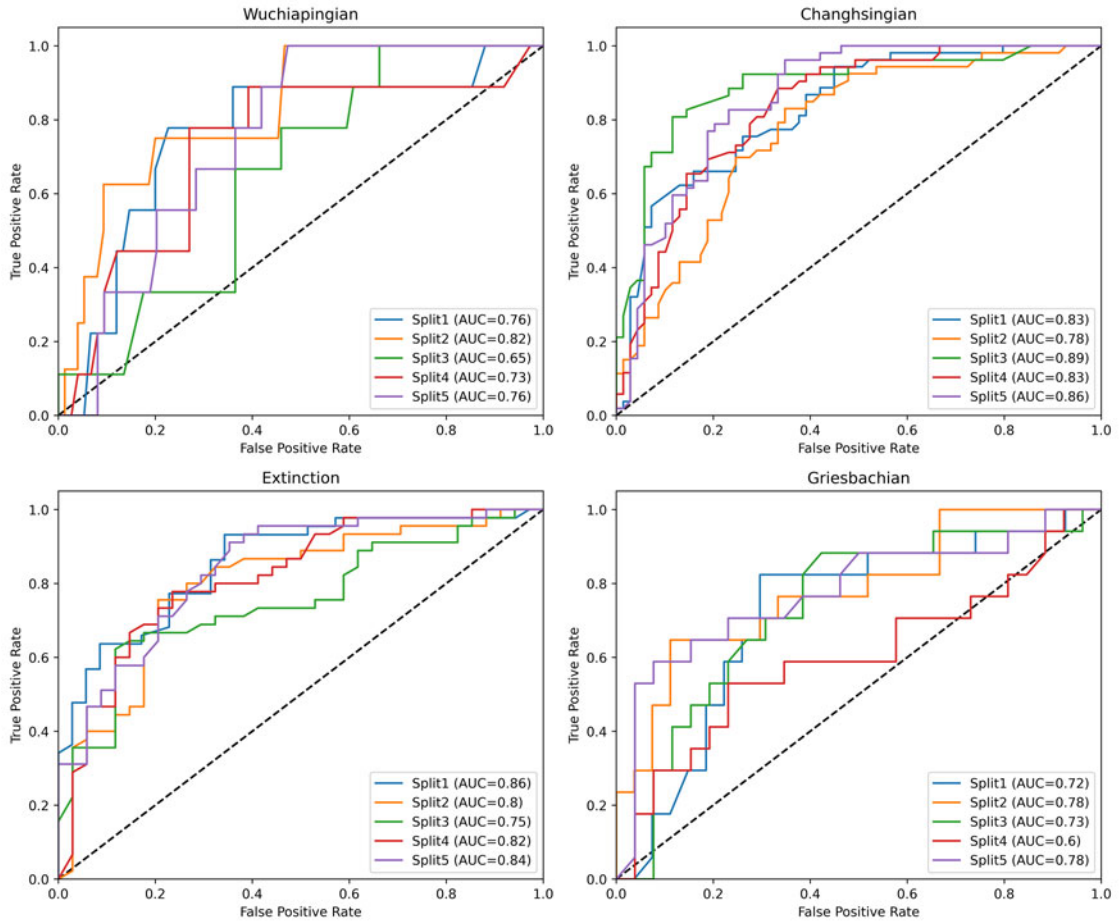
FIGURE 3. Area under the receiver operating characteristic (AUC) curves for the splits used for cross-validation in each time interval. This curve plots the true-positive rate against the false-positive rate. The black dashed line represents an AUC of 0.5, which indicates a model with a random classification that has no utility. It is unlikely that a decision tree algorithm will give an AUC value of 1, which represents a perfect classification; instead, an AUC > 0.7 is typically considered representative of a good model (e.g., the splits for the pre-extinction Changhsingian and the extinction interval).

In this study, positive SHAP values indicate that the respective feature is a good predictor of extinction, and negative values indicate a good predictor of survival.

The end-Permian mass extinction was highly selective according to an organism´s mineralogy, with the SHAP summary plot indicating that CatBoost would predict a lower extinction risk for genera with no shell or an organophosphatic shell (i.e., large negative SHAP values in Fig. 5) and a higher extinction risk for genera with a siliceous shell (i.e., large positive SHAP values in Fig. 5). This can also be seen in the SHAP value attributes for individual genera, wherein mineralogy for these mineralogical groups has one of the greatest importances in model predictions (Fig. 6A–C). The selectivity of the other mineralogical groups is less clear, because the SHAP values are near 0 and are less important for other mineralogical groups (e.g., the low-Mg calcite brachiopods; Fig. 6D, E). Aragonite, high-Mg calcite, and low-Mg calcite genera are more likely to go extinct than the remaining mineralogical groups, but this only has a small influence on CatBoost (Fig. 5), owing to the fact that the extinction percentage based on these three functional traits is between 50% and 60% (i.e., random).

Physiology and motility are also important for the CatBoost model during the extinction
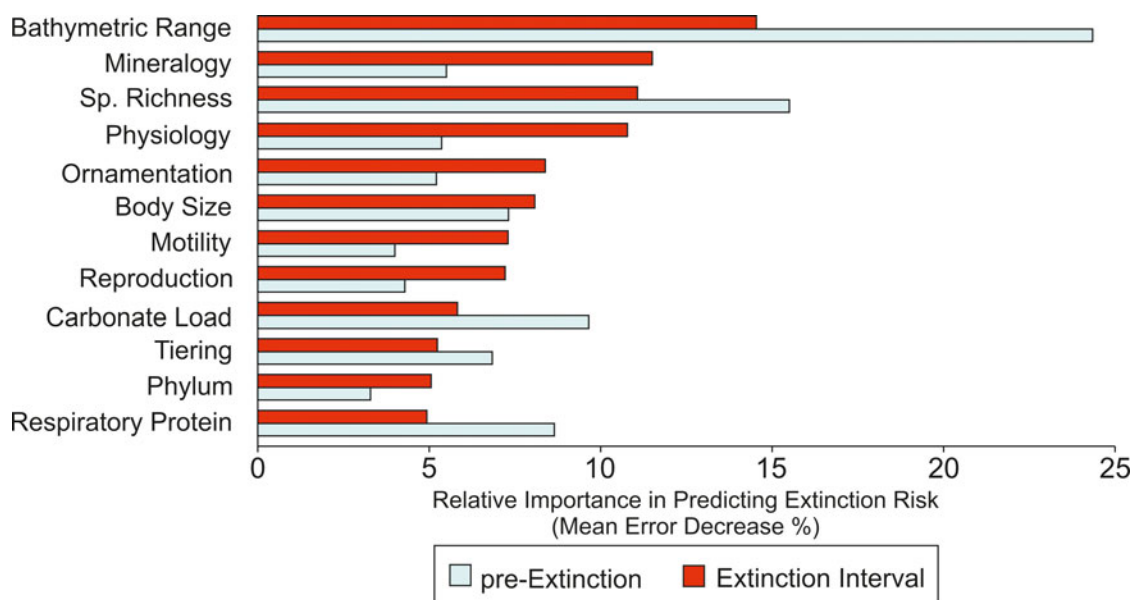
FIGURE 4.   The relative importance of each ecological variable for identifying extinction selectivity based on marine genera from South China.

interval (Fig. 5). When considering physiology, the high SHAP values show that genera with a heavy carbonate load and no buffering capacity were selected against, whereas animals with a moderate carbonate load and a moderate buffering capacity were less likely to go extinct (Fig. 5). Genera with little or no carbonate load, however, show a mixed signal, because this group includes siliceous taxa that were highly selected against and taxa with no shell that were not selected against. This means that for non-siliceous genera, for example, the articulate brachiopod *Costatumulus* (Fig. 6D), physiology is a good predictor of extinction risk, whereas for siliceous taxa, mineralogy is a better predictor of extinction risk (Fig. 6A, B). When looking at the extinction magnitude between the different types of motility, genera that are motile are less likely to have gone extinct than stationary genera (Fig. 5). The Cat-Boost also shows that for facultatively motile taxa, motility is the most important factor for predicting a low extinction risk (Fig. 5).

Discussion

The number of species within a genus was a good predictor of extinction for the Changhsingian and extinction interval, where species-rich genera were more likely to survive and species-poor genera were more likely to go extinct. An important role for species richness in the end-Permian mass extinction event has previously been reported but was overlooked, as other factors were interpreted as more important (Payne and Finnegan 2007; Clapham and Payne 2011). The importance of the species richness of a genus can be interpreted as a biological signal; it would be generally expected that genera with a high number of species have a larger pool of variation and collectively a greater tolerance to a wide range of environmental conditions. There will, however, be exceptions, as some species will have wider environmental tolerances than multiple species combined (e.g., Farrell 2009; Verberk et al. 2016). The number of species also explains some of the within-phylum selectivity; for example, articulate brachiopod genera with a higher number of species were more likely to survive (compare Fig. 6D with Fig. 6E).

The extinction selectivity according to an organism's bathymetric range was the best predictor of extinction for both the pre-extinction and extinction intervals (Fig. 4). The importance of this feature, however, decreases in the
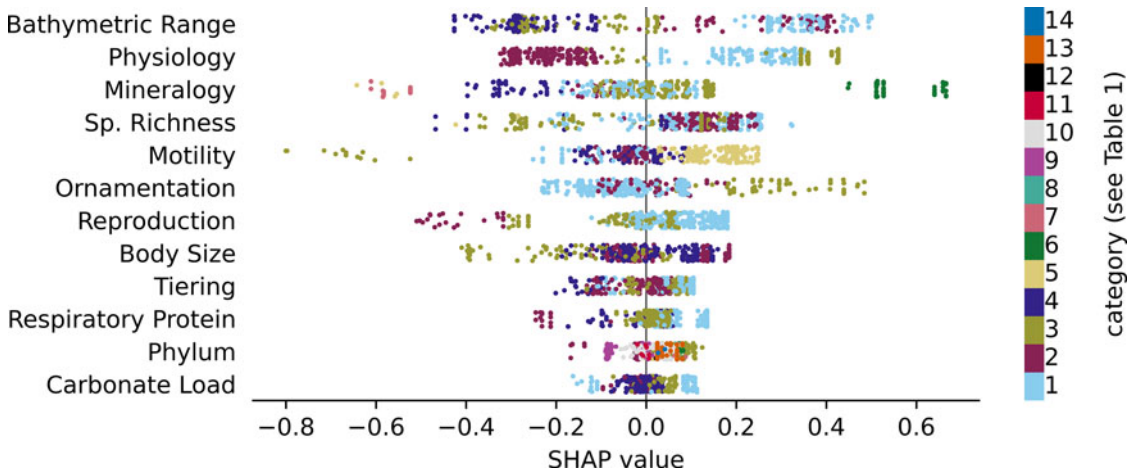
FIGURE 5. Shapley additive explanations (SHAP) summary plot showing how the different values of each ecological attribute affect the model predictions for the extinction interval. The horizontal location of the values shows whether a data point from the training dataset is associated with a higher or lower prediction. The vertical position corresponds to the relative importance of each ecological attribute. The SHAP summary plot for split 5 is shown, because of its high AUC value and because it resembles the average for all the splits combined. SHAP plots for all splits are available in Supplementary Figs. S8 and S9. The points are colored according to the categorical value given in Table 1 for each feature. For example, the feature "Mineralogy" has seven values (1, aragonite; 2, high-/intermediate-Mg calcite; 3, low-Mg calcite; 4, bimineralic; 5, organophosphatic; 6, silica; 7, no shell) and uses the first seven colors of the palette.

extinction interval. Broad bathymetric ranges are expected to enhance survivorship, because genera are expected to have tolerated a broad range of environmental conditions, but this signal is expected to be less strong during mass extinction events (Jablonksi 1986), which is also supported by our results. Of the organisms with restricted bathymetric ranges, the extinction percentage of those that were restricted to basinal settings increases from 38% to 86% during the extinction interval, whereas the extinction rate of organisms restricted to platform settings only changes from 65% to 68%. As indicated by redox proxies from South China (Song et al. 2012), this is tied to the expansion of oxygen minimum zones and the concomitant reduction of habitable areas in basinal settings (He et al. 2015). The loss of these genera from basinal settings is not considered a consequence of a sea-level change, sampling bias, or rock record bias, because postextinction basinal facies in South China have also been intensely studied but do not yield many body fossils. Shoaling of the carbonate compensation depth (CCD), which is the depth at which carbonate is undersaturated, leading to dissolution of marine shells, may also explain the selective loss of genera in basinal settings, as increases in dissolved CO_2 will lead to calcium carbonate undersaturation. Notably, some basinal successions became devoid of carbonate in the upper Changhsingian (He et al. 2005), which has been inferred as deposition below the CCD (e.g., Isozaki 1997). Understanding the processes that govern the shoaling and stability of the CCD during the Paleozoic is, however, speculative because changes in the position of the CCD in modern oceans involves important pelagic calcifiers, that is, planktic foraminifera and coccolithophores, that did not become an important component of marine ecosystems until the mid-Mesozoic. During the extinction interval, shallow platform settings (below fair-weather wave base) were also exposed to transient anoxic conditions (Song et al. 2012), which may explain the high extinction risk in this setting. The selective loss of genera restricted to deep-water habitats suggests that the expansion of the oxygen minimum zone in basinal settings was one of the main drivers of extinction.

The end-Permian extinction was highly selective against siliceous taxa compared with calcareous genera (Fig. 5). A drop in the diversity and abundance of both radiolarians and silicisponges led to a crash in biosiliceous
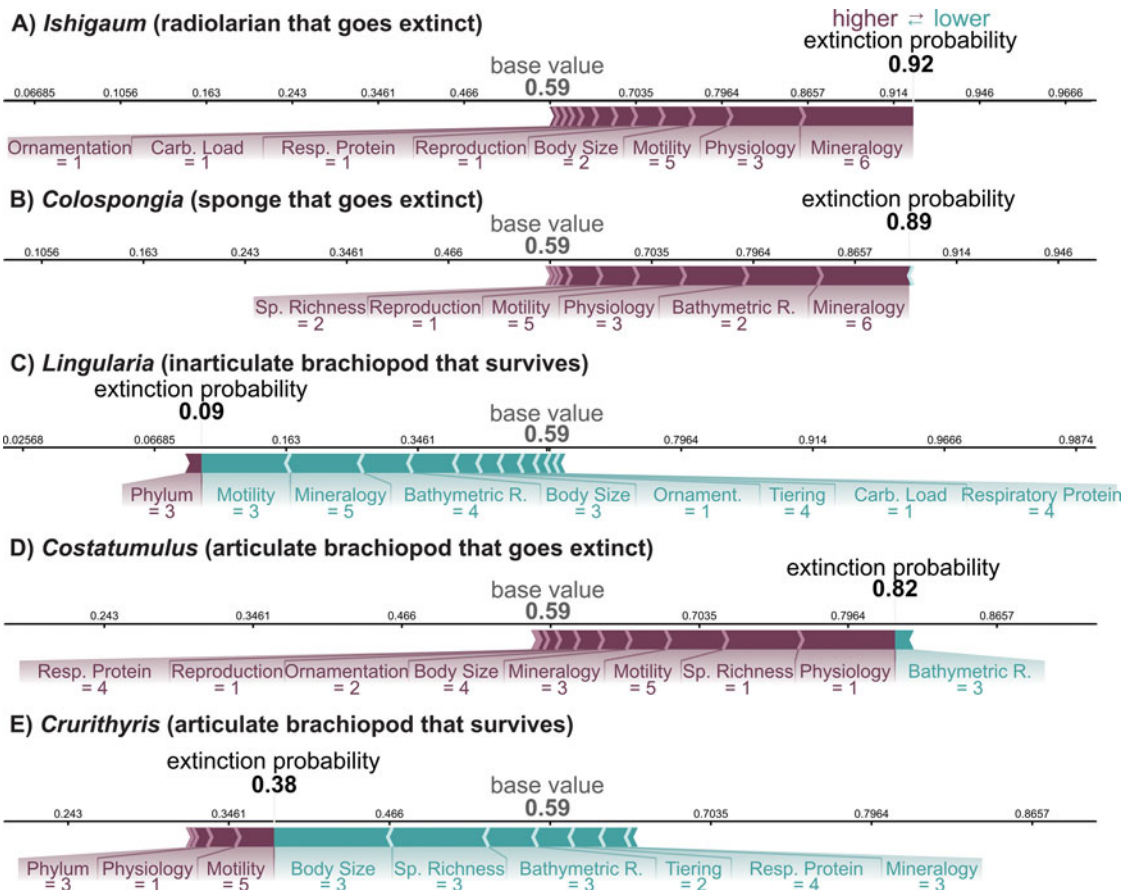
FIGURE 6. Relative Shapley additive explanations (SHAP) value attributes showing how the different ecological variables for five example genera—A, *Ishigaum*, B, *Coelocladiella*, C, *Lingularia*, D, *Costatumulus*, and E, *Crurithyris*—change in the model prediction for the extinction interval. The x-axis is the model output value. The base value is the prediction if no ecological variables are considered, that is, the average predicted probability. Model output value is the prediction considering the ecological variables for the investigated genus, with positive values indicating extinction and negative values indicating survival. Features that push the prediction higher (to the right), that is, more likely to go extinct, are shown in burgundy, and the opposite is shown in blue. The categorical values for the functional traits are given in Table 1 for each feature.

ooze production, resulting in an equatorial chert gap during the Early Triassic (Racki 1999). One factor that could have led to the demise of effective silica factories during the end-Permian extinction is rapidly rising ocean temperatures that would have increased silica dissolution rates and altered the silica saturation state (Racki 1999; Beauchamp and Baud 2002). Temperature stress is also expected to produce selective patterns of extinction by preferentially extinguishing genera that already lived near their upper thermal limits (Song et al. 2014). Based on what is known on the thermal limits of modern radiolarians (Song et al. 2014), the temperatures that developed

during the extinction interval in South China (Joachimski et al. 2012) likely made equatorial surface waters uninhabitable. Further support that high temperatures caused this selective signal is that radiolarians, silicisponges, and chert deposits that do occur after the mass extinction event are only known from thermal refugia, that is, high paleolatitudes or deeper-water settings (Takemura et al. 2002; Godbold et al. 2017), and that in shallow-marine settings, the major metabolic stress in the tropics was thermal stress (Penn et al. 2019).

The preferential extinction of siliceous genera does not negate ocean acidification and instead highlights that rapid increase in CO_2 into the

ocean–atmosphere system has numerous complex consequences on ecosystems. Experimental studies investigating the impact of ocean acidification on the biological calcification of marine animals have shown that species that precipitate aragonite and high-Mg calcite are more vulnerable than those that precipitate low-Mg calcite (Ries 2011). Even though our results show that genera that precipitate siliceous skeletons were more susceptible to extinction than any other mineralogical group, aragonite, high-Mg calcite, and low-Mg calcite genera were also significantly selected against when compared with genera that possess no shell (Fig. 5). Despite this selectivity pattern that is consistent with an ocean acidification scenario, the slightly higher extinction likelihood of genera with low-Mg calcite skeletons over aragonite skeletons during the extinction interval is inconsistent with modern studies (e.g., Ries et al. 2009; Ries 2011) suggesting that ocean acidification was not a selective pressure during the extinction interval.

It is not only a genus' CaCO$_3$ polymorph that dictates whether it is vulnerable to ocean acidification. For instance, an organism's ability to regulate pH and carbonate chemistry at the site of calcification, the degree to which the organism's biomineral is protected by an organic coating, and shell microstructure are characteristics that have also been noted (Ries et al. 2009; Ries 2011; Garbelli et al. 2017). Even though these factors are important, it is not yet possible to include these in a detailed deep-time extinction selectivity analysis, as this information is largely unknown, even for extant organisms. Physiological classifications to include these factors have been performed, but only at the class level (Knoll et al. 2007). These studies found a selectivity pattern consistent with an ocean acidification event. Using the same physiological categories, our study found a selective pattern that is inconsistent with ocean acidification and previous studies; that is, genera with little or no carbonate load were more strongly selected against than "buffered" genera with calcareous skeletons (Fig. 5), owing to the selective loss of siliceous genera. When siliceous taxa are removed from the study, however, the extinction pattern is consistent with an ocean acidification scenario

(Fig. 5). Our study also shows that physiology was one of the most important ecological factors for predicting extinction in non-siliceous taxa (e.g., *Costatumulus*; Fig. 6D). Notably, the few studies that have been able to compare some of these additional factors within phyla have found a selectivity pattern consistent with ocean acidification, for example, differences in the shell microstructure of Strophomenata and Rhynchonellata brachiopods (Garbelli et al. 2017).

The remaining variable that had a selective signal on extinction risk is motility. Genera that are motile are less likely to have gone extinct than stationary genera (Fig. 5), consistent with previous observations (Foster and Twitchett 2014; Clapham 2017). The resilience of motile genera to rapid climate warming is likely due to motile genera typically having a greater aerobic scope, inherently high extracellular pCO$_2$, and higher maximum thermal tolerance limits (Clapham 2017). The selective loss of stationary genera did not exist before the mass extinction, which suggests that conditions responsible for this selective pattern developed rapidly during the extinction interval. The preferential extinction of stationary genera compared with motile taxa cannot be used to infer individual extinction mechanisms, as the physiological adaptations of motile taxa are advantageous under high-temperature, low-oxygen, and low-pH conditions.

Recently, the biogeographic and physiological selectivity of the end-Permian mass extinction has been interpreted to have been a consequence of a combination of high temperatures and widespread marine anoxia (Penn et al. 2019). Our CatBoost results, however, show that ecological selectivity during the end-Permian extinction varies among different groups of organisms and that a synergy of multiple environmental stressors best explains the overall end-Permian extinction selectivity patterns. Given that the selectivity patterns observed for the extinction interval in this study are interpreted as a consequence of expanded oxygen minimum zones, high sea-surface temperatures, and ocean acidification, this selectivity pattern corroborates the inferences of previous studies (i.e., Knoll et al. 2007; Clapham and Payne 2011) that suggest

multiple factors synergistically drove the end-Permian mass extinction event and that this "deadly trio of carbon dioxide" is responsible for the ecological selectivity signal at the end-Permian extinction.

Conclusions

Decision tree algorithms have a great potential to reveal selectivity patterns during both past and projected extinction events. Applying the categorical gradient boosting algorithm to a database of marine invertebrates, conodonts, and calcareous algae that span the Permian/Triassic boundary in South China, our analysis reveals that extinction risk was greater for genera that had a low species richness, were limited to deep-water habitats, had a stationary mode of life, possessed a siliceous skeleton, or, less critically, had calcitic skeletons. Furthermore, these selectivity patterns did not exist before the extinction, suggesting that extinction drivers changed between the pre-extinction interval and the extinction interval. We linked these selective losses to the synergistic effects of rapid climate change associated with the end-Permian mass extinction, that is, expanded oxygen minimum zones, rapid ocean warming, and ocean acidification.

Acknowledgments

Data Availability Statement

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.hmgqnk9j7.

Literature Cited

Aberhan, M., W. Kiessling, and F. T. Fürsich. 2006. Testing the role of biological interactions in the evolution of mid-Mesozoic marine benthic ecosystems. Paleobiology 32:259–277.

Ayaru, L., P. P. Ypsilantis, A. Nanapragasam, R. C. H Choi, A. Thillanathan, L. Min-Ho, and G. Montana. 2015. Prediction of outcome in acute lower gastrointestinal bleeding using gradient boosting. PLoS ONE 10:e0132485.

Bambach, R. K., A. M. Bush, and D. H. Erwin. 2007. Autoecology and the filling of ecospace: key metazoan radiations. Palaeontology 50:1–22.

Beauchamp, B., and A. Baud. 2002. Growth and demise of Permian biogenic chert along northwest Pangea: evidence for end-Permian collapse of thermohaline circulation. Palaeogeography, Palaeoclimatology, Palaeoecology 184:37–63.

Belyaeva, A., S. Venkatachalapathy, M. Nagarajan, G. V. Shivashanker, and C. Uhler. 2017. Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription. Proceedings of the National Academy of Sciences USA 1144:13714–13719.

Boyer, A. G. 2010. Consistent ecological selectivity through time in Pacific Island avian extinctions. Conservation Biology 24:511–519.

Breiman, L., J. Friedman, C. J. Stone, and R. A. Olshen. 1984. Classification and regression trees. Routledge, Boca Raton, Fla.

Burgess, S. D., and S. A. Bowring. 2015. High-precision geochronology confirms voluminous magmatism before, during, and after Earth's most severe extinction. Science Advances 1:e1500470.

Bush, A. M., G. Hunt, and R. K. Bambach. 2016. Sex and the shifting biodiversity dynamics of marine animals in deep time. Proceedings of the National Academy of Sciences USA 113:14073–14078.

Bzdok, D., N. Altman, and M. Krzywinski. 2018. Statistics versus machine learning. Nature Methods 15:233–234.

Cawley, G. C., and N. L. Talbot. 2010. On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of Machine Learning Research 11:2079–2107.

Clapham, M. E. 2017. Organism activity levels predict marine invertebrate survival during ancient global change extinctions. Global Change Biology 23:1477–1485.

Clapham, M. E., and J. L. Payne. 2011. Acidification, anoxia, and extinction: a multiple logistic regression analysis of extinction selectivity during the Middle and Late Permian. Geology 39:1059–1062.

Dorogush, A. V., V. Ershov, and A. Gulin. 2018. CatBoost: gradient boosting with categorical features support. arXiv preprint 1810.11363.

Dunhill, A. M., W. J. Foster, S. Azaele, J. Sciberras, and R. J. Twitchett. 2018. Modelling determinants of extinction across two Mesozoic hyperthermal events. Proceedings of the Royal Society of London B 285:20180404.

Farrell, A. P. 2009. Environment, antecedents and climate change: lessons from the study of temperature physiology and river migration of salmonids. Journal of Experimental Biology 212:3771–3780.

Finnegan, S., N. A. Heim, S. E. Peters, and W. W. Fischer. 2012. Climate change and the selective signature of the Late Ordovician mass extinction. Proceedings of the National Academy of Sciences USA 109:6829–6834.

Finnegan, S., S. C. Anderson, P. G. Harnick, C. Simpson, D. P. Tittensor, J. E. Byrnes, Z. V. Finkel, D. R. Lindberg, L. H. Liow, R. Lockwood, H. K. Lotze, C. R. McClain, J.

L McGuire, A. O'Dea, and J. M. Pandolfi. 2015. Paleontological baselines for evaluating extinction risk in the oceans. Science 348:567–570.

Foote, M. 2000. Origination and extinction components of taxonomic diversity: general problems. Paleobiology 26:74–102.

Foster, W. J., and R. J. Twitchett. 2014. Functional diversity of marine ecosystems after the Late Permian mass extinction event. Nature Geoscience 7:233–238.

Friedman, J. H. 2001. Greedy function approximation: a gradient boosting machine. Annals of Statistics 29:1189–1232.

Garbelli, C., L. Angiolini, and S-Z. Shen. 2017. Biomineralization and global change: a new perspective for understanding the end-Permian extinction. Geology 45:19–22.

Godbold, A., S. Schoepfer, S-Z. Shen, and C. M. Henderson. 2017. Precarious ephemeral refugia during the earliest Triassic. Geology 45:607–610.

Hanna, E., and M. Cardillo. 2014. Island mammal extinctions are determined by interactive effects of life history, island biogeography and mesopredator suppression. Global Ecology and Biogeography 23:395–404.

He, W. H., S.-Z. Shen, Q. L. Feng, and S. A. Gu. 2005. Late Changhsingian (Late Permian) deepwater brachiopod fauna from the Talung Formation at the Dongpan section, Southern Guangxi, South China. Journal of Paleontology 79:927–938.

He, W. H., G. R. Shi, R. J. Twitchett, Y. Zhang, K. X. Zhang, H. J. Song, M. L. Yue, S. B. Wu, H. T. Wu, T. L. Yang, and Y. F. Xiao. 2015. Late Permian marine ecosystem collapse began in deeper waters: evidence from brachiopod diversity and body size changes. Geobiology 13:123–138.

Isozaki, Y. 1997. Permo-Triassic boundary superanoxia and stratified superocean: records from lost deep sea. Science 276:235–238.

Jablonski, D. 1986. Background and mass extinctions: the alternation of macroevolutionary regimes. Science 231:329–133.

Jin, Y. G., Y. Wang, W. Wang, Q. H. Shang, C. Q. Cao, and D. H. Erwin. 2000. Pattern of marine mass extinction near the Permian–Triassic boundary in South China. Science 289:432–436.

Joachimski, M. M., X. Lai, S. Shen, H. Jiang, G. Luo, B. Chen, J. Chen, and Y. Sun. 2012. Climate warming in the latest Permian and the Permian–Triassic mass extinction. Geology 40:195–198.

Kaggle. 2022. Titanic—machine learning from disaster. https://www.kaggle.com/c/titanic.

Karapatne, A., I. Ebert-Uphoff, S. Ravela, H. Ali Babaie, and V. Kumar. 2017. Machine learning for the geosciences: challenges and opportunities. IEEE Transactions on Knowledge and Data Engineering 31:1–12.

Knoll, A. H., R. K. Bambach, J. L. Payne, S. Pruss, and W. W. Fischer. 2007. Paleophysiology and end-Permian mass extinction. Earth and Planetary Science Letters 256:295–313.

Kocsis, Á. T., C. J. Reddin, and W. Kiessling. 2018. The biogeographical imprint of mass extinctions. Proceedings of the Royal Society of London B 285:20180232.

Lundberg, S. M., and S.-I. Lee. 2017. A unified approach to interpreting model predictions. Pp. 4765–4774 in NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Curran Associates, Red Hook, N.Y.

Mohri, M., A. Rostamizadeh, and A. Talwalkar. 2018. Foundations of machine learning, 2nd ed. MIT Press, Cambridge, Mass.

Natekin, A, and A. Knoll. 2013. Gradient boosting machines, a tutorial. Frontiers in Neurorobotics 27:1–21.

Novack-Gottshall, P. M. 2016. General models of ecological diversification. II. Simulations and empirical applications. Paleobiology 42:209–239.

Payne, J. L., and S. Finnegan. 2007. The effect of geographic range on extinction risk during background and mass extinction. Proceedings of the National Academy of Sciences USA 104:10506–10511.

Payne, J. L., D. J. Lehrmann, D. Follett, M. Seibel, L. R. Kump, A. Riccardi, D. Altiner, H. Sano, and J. Wei. 2007. Erosional truncation of uppermost Permian shallow-marine carbonates and implications for Permian–Triassic boundary events. Geological Society of America Bulletin 119:771–784.

Payne, J. L., A. M. Bush, E. T. Chang, N. A. Heim, M. L. Knope, and S. B. Pruss. 2016. Extinction intensity, selectivity and their combined macroevolutionary influence in the fossil record. Biology Letters 12:e20160202.

Penn, J. L., C. Deutsch, J. L. Payne, and E. A. Sperling. 2019. Temperature-dependent hypoxia explains biogeography and severity of end-Permian marine mass extinction. Science 362:1130–1136.

Prokhorenkova, L., G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin. 2018. CatBoost: unbiased boosting with categorical features. Pp. 6639–6649 in Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada.

Racki, G. 1999. Silica-secreting biota and mass extinctions: survival patterns and processes. Palaeogeography, Palaeoclimatology, Palaeoecology 154:107–132.

Ries, J. B. 2011. Physicochemical framework for interpreting the biological calcification response to CO_2-induced ocean acidification. Geochimica et Cosmochimica Acta 75:4053–4064.

Ries, J. B., A. L. Cohen, and D. C. McCorkle. 2009. Marine calcifiers exhibit mixed responses to CO_2-induced ocean acidification. Geology 37:1131–1134.

Scotese, C. R. 2016. PALEOMAP PaleoAtlas for GPlates and the PaleoData Plotter Program. https://www.earthbyte.org/paleo-map-paleoatlas-for-gplates, accessed 7 January 2018.

Smith, S. M., K. D. Angielczyk, L. Schmitz, and S. C. Wang. 2018. Do bony orbit dimensions predict diel activity pattern in Sciurid rodents? Anatomical Record 301:1774–1787.

Song, H., P. B. Wignall, J. Tong, D. P. Bond, H. Song, X. Lai, K. Zhang, H. Wang, and Y. Chen. 2012. Geochemical evidence from bio-apatite for multiple oceanic anoxic events during Permian–Triassic transition and the link with end-Permian extinction and recovery. Earth and Planetary Science Letters 353–354:12–21.

Song, H., P. B. Wignall, J. Tong, and H. Yin. 2013. Two pulses of extinction during the Permian–Triassic crisis. Nature Geoscience 6:52–56.

Song, H., P. B. Wignall, D. Chu, J. Tong, Y. Sun, H. Song, W. He, and L. Tian. 2014. Anoxia/high temperature double whammy during the Permian–Triassic marine crisis and its aftermath. Scientific Reports 4:4132.

Svensen, H., S. Planke, A. G. Polozov, N. Schmidbauer, F. Corfu, Y. Y. Podladchikov, and B. Jamtveit. 2009. Siberian gas venting and the end-Permian environmental crisis. Earth and Planetary Science Letters 277:490–500.

Takemura, A., Y. Aita, R. S. Hori, Y. Higuchi, K. B. Spörli, H. J. Campbell, K. Kodama, and T. Sakai. 2002. Triassic radiolarians from the ocean-floor sequence of the Waipapa Terrane at Arrow Rocks, Northland, New Zealand. New Zealand Journal of Geology and Geophysics 45:289–296.

Tietje, M., and M. O. Rödel. 2018. Evaluating the predicted extinction risk of living amphibian species with the fossil record. Ecology Letters 21:1135–1142.

Verberk, W. C. E. P, J. Overgaard, R. Ern, M. Bayley, T. Wang, L. Boardman, and J. S. Terblanche. 2016. Does oxygen limit thermal tolerance in arthropods? A critical review of current evidence. Comparative Biochemistry and Physiology Part A: Molecular and Integrative Physiology 192:64–78.

Wang, Y., P. M. Sadler, S. Z. Shen, D. H. Erwin, Y. C. Zhang, X. D. Wang, W. Wang, J. L. Crowley, and C. M. Henderson. 2014. Quantifying the process and abruptness of the end-Permian mass extinction. Paleobiology 40:113–129.

Wignall, P. B., and R. J. Twitchett. 1996. Oceanic anoxia and the end-Permian mass extinction. Science 272:1155–1158.

Yuan, D. X., S. Z. Shen, C. M. Henderson, J. Chen, H. Zhang, and H. Z. Feng. 2014. Revised conodont-based integrated high-resolution timescale for the Changhsingian Stage and end-Permian extinction interval at the Meishan sections, South China. Lithos 204:220–245.