Project Proposal

Yutao ZHOU

Affiliation: Imperial College London

Email: lz723@ic.ac.uk

Supervisor

Sammraat Pawar, Imperial College London

Date: 30 Mar 2024

1   Introduction

In the pursuit of understanding how organisms' physiological rates are influ-
5   enced by temperature, Thermal Performance Curves (TPCs) have emerged as
a cornerstone concept in ecological and evolutionary biology Kontopoulos et al.
(2020)(more article). These curves, which depict the relationship between tem-
perature and biological rates, are integral for predicting the impacts of climate
change on biological systems from individual organisms to entire ecosystems
10 Mordecai et al. (2019)(more). The complexity and diversity of TPC models
necessitate robust computational tools for their analysis. Padfield, O'Sullivan,
and Pawar (2021) introduced a significant advancement in this arena with the
development of the rTPC and nls.multstart R packages. This innovative com-
putational pipeline facilitates the fitting of TPCs to large datasets, offering a
15 selection from 24 distinct TPC models and addressing the urgent need for scal-
able and versatile tools in the face of burgeoning ecological and evolutionary
data.

However, the question of optimal model selection remains a formidable chal-
20 lenge, given the array of available models and their varying performance across
different datasets. Kontopoulos, Sentis, Daufresne, Dell, and Pawar (2021) made
strides in addressing this challenge by conducting a systematic comparison of
83 TPC models across a broad spectrum of traits and taxonomic groups. Their
study highlighted the considerable variability in model performance and under-
25 scored the absence of a one-size-fits-all model for TPC analysis. This finding
emphasizes the necessity of a thoughtful model selection process to mitigate
potential biases and enhance the accuracy of TPC analyses. Integrating the
insights from Kontopoulos et al. (2021) into the rTPC pipeline initiated by
Padfield et al. (2021) offers a promising avenue for refining and expanding the
30 toolkit available to researchers, thus paving the way for a deeper and more nu-
anced understanding of organisms' responses to temperature changes within the
context of global climate dynamics.

2   Proposed Methods

2.1   Data Prepossessing

35 Utilizing tools from the rTPC package, we first standardize the preprocessing of
ecological and evolutionary biology datasets, including adjustments to temper-
ature ranges and outlier processing. This step ensures consistency across data,
laying a solid foundation for subsequent analyses.

2.2 Model Library Construction and Selection

Drawing on the research by Kontopoulos et al. (2021), we will build a more extensive library of TPC models, incorporating not only the 24 models found in rTPC but also other models validated within ecology and evolutionary biology. From this library, we develop a machine learning-based model selection algorithm capable of dynamically recommending the most suitable model based on the characteristics and taxonomic groups of the dataset.

2.3 Adaptive Learning Algorithm

We propose an adaptive learning algorithm that continually refines model recommendations using new data sources. By analyzing the performance of each model fitting, the algorithm updates the model's weights and preferences to achieve adaptive fitting to complex ecological data.
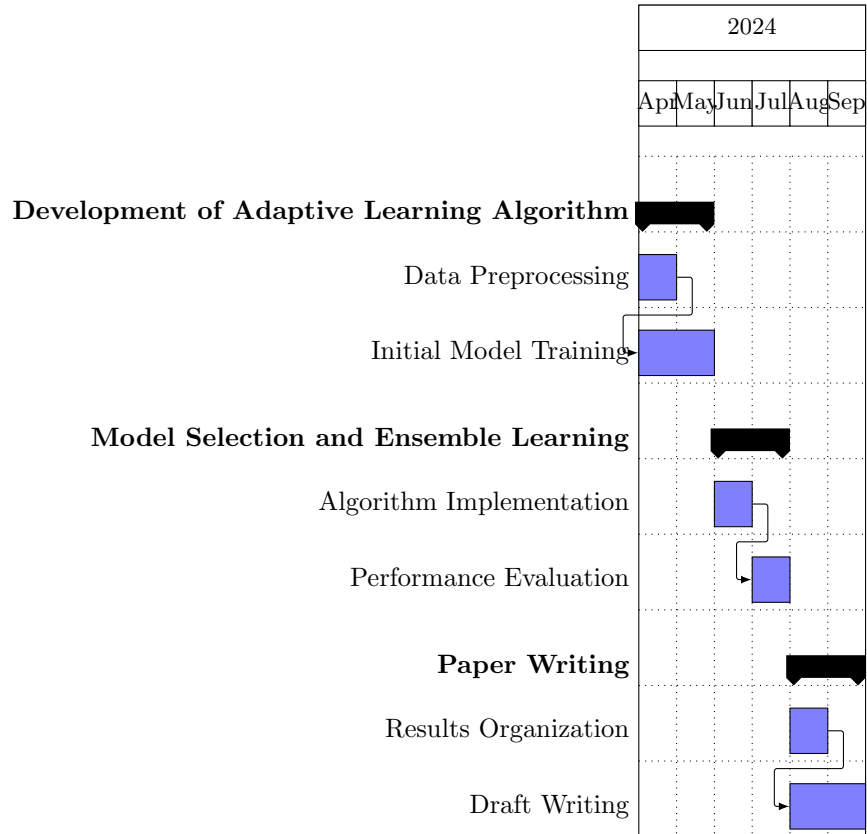
2.4 Ensemble Learning and Model Integration

An ensemble learning strategy is adopted, not to singularly select the "best" model but to integrate predictions from multiple models. Weights based on their historical performance enhance the overall prediction's accuracy and robustness.

2.5 Model Validation and Application

Statistical methods, including but not limited to cross-validation, AIC, and BIC, will be employed for model validation. Furthermore, we plan to demonstrate the application of this method in addressing real-world ecological and evolutionary problems through a series of case studies, such as predicting the impact of climate change on species distributions.

3 Timeline

	2024					
	Apr	May	Jun	Jul	Aug	Sep
Development of Adaptive Learning Algorithm						
Data Preprocessing						
Initial Model Training						
Model Selection and Ensemble Learning						
Algorithm Implementation						
Performance Evaluation						
Paper Writing						
Results Organization						
Draft Writing						

4 Budget

no budget

References

Kontopoulos, D.-G., Sentis, A., Daufresne, M., Dell, A. I., & Pawar, S. (2021). No model to rule them all: a systematic comparison of 83 thermal performance curve models across traits and taxonomic groups. *Title of the journal*, *Volume number*(Issue number), Page range.

Kontopoulos, D.-G., Van Sebille, E., Lange, M., Yvon-Durocher, G., Barraclough, T. G., & Pawar, S. (2020). Phytoplankton thermal responses adapt in the absence of hard thermodynamic constraints. *Evolution*, *74*(4), 775–790. Retrieved from `https://doi.org/10.1111/evo.13946` doi: 10.1111/evo.13946

Mordecai, E. A., Caldwell, J. M., Grossman, M. K., Lippi, C. A., Johnson, L. R., Neira, M., ... Villena, O. (2019). Thermal biology of mosquito-borne disease. *Ecology Letters*, *22*(10), 1690–1708. Retrieved from `https://doi.org/10.1111/ele.13335` doi: 10.1111/ele.13335

Padfield, D., O'Sullivan, H., & Pawar, S. (2021). rtpc and nls.multstart: A new pipeline to fit thermal performance curves in r. *Methods in Ecology and Evolution*, *12*(1138-1143), 1138–1143. doi: 10.1111/2041-210X.13585