Data-driven parameterization for climate modeling (QBO)

Mentor: Prof. Edwin Gerber
Applicant: Zihan Shao

Abstract

Quasi-Biennial Oscillation, QBO in short, is a quasiperiodic oscillation of the equatorial zonal wind between easterlies and westerlies in the tropical stratosphere with a mean period of 28 to 29 months.[1] Atmospheric Gravity Waves, GWs in short, plays an important role in the formation of QBO. This project will be focusing on the data-driven parameterization for the GWs in the QBO climate model, aiming to improve climate model's ability to simulate this vacillation of jets in the tropical stratosphere. In specific, we will focus on the 1-dimensional version of the problem.

Research Question and Significance

Atmospheric Gravity Waves(GWs), play an important role in the exchange of momentum between the Earth's surface and the free atmosphere. They slow down the tropospheric jet streams and stratospheric polar night jets, and are a key driver of the QBO. GWs present a challenge to climate prediction because they cannot be properly resolved with available computational power. They must rather be estimated from the larger scale variables that are resolved, or parameterized, as how it is described in the field. However, current physics-based gravity wave parameterizations neglect transient effects and account only for vertical wave propagation.[3] By using data-driven model, parametrizations will be developed to more accurately and efficiently represent GW momentum fluxes in weather and climate models. This, ultimately, will enhance our ability to predict the changes in climate and weather.

In this project, we will try to do the data-driven parametrization for GWs in QBO models. Followed is a mathematical characterization of the problem. In 1-dimensional version of the problem, the unknown zonal wind u, which we expect to oscillates biannually, is governed by the following equation

$$\frac{\partial u}{\partial t} + w\frac{\partial u}{\partial z} - K\frac{\partial^2 u}{\partial z^2} = -S(u, z, A, c)$$

Here u is the unknown zonal wind, t is the time, z is the vertical coordinate, w and K are constant vertical advection and diffusivity, and $S(u, z, A, c)$ is the wave forcing consisting of a sum of monochromatic waves, which are the Gravity Waves we mentioned above[2]. Note that here u is internally generated by the model and A and c are externally forced (that is, supplied as some external condition), which we take to be generated by a random variable. For now, we have a physical parametrization of S. This, together with the PDE above, will generate a series of (S, u, z, A, c), which will be treated as the ground truth.

In that case, we are going to design a Machine Learning model to predict S given (S, u, z, A, c) generated by the physical parametrization. The parameters we might use includes but not restricted to: u(zonal wind), z(vertical coordinates), A(amplitude) and c(phase speed).

Note that the evaluation metric for this problem will be different from the usual Data Science problem. On the one hand, the predicted S should be close to the ground truth. On the other hand, the predicted S, together with the PDE, should generate the correct oscillation. In general, the second criterion is more important, since we want the predicted model to be used in actual climate prediction.

Project Design and Feasibility

Our plan is to divide the project into the following two parts.

1. The preliminaries. Before the actual Machine Learning task, we need to figure out what exactly QBO is. This includes, the physical formation of QBO, and how the physical parametrization of the Gravity waves works.

2. Machine Learning tasks. After the acquisition of the preliminaries, we could start doing the Machine Learning tasks. To be specific, we are going to implement some basic machine learning algorithms (e.g. Linear Regression) and then apply some more advanced algorithms and even state-of-the-art Deep Learning algorithms.

I'm going to work at least 3 hour a day to make sure I can complete the project by the end of the summer. The meeting will be conducted online (via Zoom).

Background

I'm a junior student majoring in Honors Mathematics and Data Science. At the current stage, I have taken all foundational math courses including Honors Calculus, Honors Analysis I and II, Honors Linear Algebra I and II, Honors Theory of Probability, ODE, Functions of Complex Variables, Real Variables. This will help me to understand the mathematical material in this project. Also, I have also taken some Data Science related courses, such as Machine Learning and Computer Vision, which equip me with the ability to do the Machine Learning part of the project.

Feedback and Evaluation

In general, I plan to hold biweekly meetings with our mentor, Professor Edwin Gerber, to update my research progress, discuss tough problems we may meet and communicate future steps. Below is our tentative schedule for the project (subject to changes according to actual research progress)

- 6.1 - 6.15. Preliminaries: Basic knowledge of QBO and Gravity Waves.
- 6.15 - 7.1. Experiments with basic Machine Learning algorithms.
- 7.1 - 7.18. Analyze the outcomes of the previous trials, and decide which algorithms we are going to focus on.
- 7.18 - 8.15. The optimization of the Machine Learning models. This also includes the trials of deep learning algorithms.
- 8.15 - 8.25. Wrap up the outcome (Including writing a paper).

Dissemination of Knowledge

The final results of our project will be presented in a paper(together with codes), summarizing the process, findings and our reflections. We are also willing to share our work with the community during any possible academic events or research symposiums.

References

[1] Baldwin, M. P. et al. The quasi-biennial oscillation. Reviews of Geophysics, 39(2), 179–229. https://doi.org/10.1029/1999rg000073

[2] A webpage of 1d QBO probelm. https://datawaveproject.github.io/qbo1d/html/model-description.html

[3] A Data-informed Framework for the Representation of Sub-grid Scale Gravity Waves to Improve Climate Prediction.