The Differentiable Ecosystem Model (a hybrid, physics-informed machine learning system for ecosystem modeling). This release contains example codes and datasets similar to those used to produce some work in the following paper:

Aboelyazeed, D., Xu, C., Hoffman, F. M., Jones, A. W., Rackauckas, C., Lawson, K. E., and Shen, C.: A differentiable ecosystem modeling framework for large-scale inverse problems: demonstration with photosynthesis simulations, Biogeosciences Discuss. [preprint], https://doi.org/10.5194/bq-2022-211, in review, 2022. I

If you have any questions for this code release, feel free to contact us by dmf5963@psu.edu (Doaa Aboelyazeed) or cshen@engr.psu.edu (Chaopeng Shen)

Follow the below instructions to run the code:

1. Set up the environment

Use the file called <u>environment.yml</u> under the main directory 'diffEcosys' to create the same conda environment.

- First : install conda
- Second: run the following command to create the environment 'conda env create -f environment.yml'
- Third : run the following command to activate the environment 'conda activate diffEcosys'

2. Check the data

An example dataset was uploaded under the directory:

'/diffEcosys/data/example_dataset/example_dataset.csv'

Please check the dataset and the attached description file that includes the required information about the name of different variables and their units.

3. Run scripts

Use scripts located under the directory '/diffEcosys/synthetic_scripts' and run them sequentially as the following:

 Run Forward_run.py first to create the synthetic observations for the data in example_dataset.csv. You need to change between: option = 0 "single parameter recovery (V_{c.max25}) only" option = 1 "Dual parameter recovery ($V_{c,max25}$ - B)" You can check the output files from this step called "Synthetic_Data_OnePAR.csv" and "Synthetic_Data_TwoPAR.csv" corresponding to options 0 and 1 respectively.

• Run **Synthetic_run.py** second to run the parameter retrieval experiments for both options