Challenge 2: **Modeling CO_2 fluxes and evapotranspiration in grain-only and graze-out wheat using weather and satellite remote sensing data.**

Accurate estimation of flux changes in agricultural ecosystems is essential for comprehending ecosystem dynamics, evaluating land productivity, and monitoring carbon and water cycles. This information guides agricultural management practices, including irrigation scheduling, crop management, and carbon sequestration strategies. However, deploying and maintaining eddy covariance systems for continuous measurement of CO_2 fluxes and evapotranspiration (ET) is not always feasible. Therefore, there is a need to explore alternative approaches, such as predictive modeling using readily available data sources.

Thus, in this study, we aim to investigate the feasibility of developing predictive models (using machine learning and deep learning frameworks) to assess CO_2 and ET dynamics by leveraging weather data and vegetation index from remote sensing sources. The dataset consists of several wheat sites managed under similar practices over multiple years, including periods of grain-only and graze-out management. We will analyze these management scenarios separately to determine if we can accurately model flux variations in the environment. Additionally, we will compare the differences in the drivers of flux changes and modeling strategies between grain-only and graze-out management practices.

Some definitions:

- **Vegetation** is represented using EVI: Enhanced Vegetation Index. Satellite remote sensing data.
- **Ecosystem fluxes**: (ET, NEE, ER, GPP).
- **ER**: Ecosystem respiration refers to total release of CO_2 from an ecosystem to the atmosphere.
- **NEE**: Net ecosystem exchange represents the net balance of CO_2 exchange between an ecosystem and the atmosphere (uptake: photosynthesis and release: respiration).
- **GPP**: Gross primary production refers to total amount of carbon uptake by plants through photosynthesis.
- **ET**: Evapotranspiration refers to the process of water evaporation from the soil surface and transpiration from plant leaves.

Attributes and units:

- ER = Ecosystem respiration (g C m-2 d-1)
- NEE = Net ecosystem exchange (g C m-2 d-1)
- GPP = Gross primary production (g C m-2 d-1)
- ET = Evapotranspiration (mm d-1)
- Tmax = maximum air temp (F)
- Tmin = minimum air temp (F)
- Tavg = average air temp (F)
- Havg = average relative humidity (%)
- Hdeg = heating degree-days (65 F standard)
- Cdeg = cooling degree-days (65 standard)
- Wcmn = minimum wind chill index temp (F)
- Wspd = average wind speed (mph)
- Atot = solar radiation (MJ/m2)
- Rain = daily rainfall (inch)

Dataset:

We have two datasets for Grain-Only and Graze-Out management practices. For each dataset:

- **Time period:** Multi-year seasonal (Mid Oct-Mid-June) information.
- **Crop & Site Info:** Wheat Crop under same management practices for several sites.
- **Weather**: Collected from infield measurements.
- **EVI**: Remote sensing data estimated using satellite observations (Landsat and Sentinel).
- **Fluxes**: Eddy covariance measurements.

<u>Study Objectives:</u>

We will conduct the same processes for both the datasets: Grain-Only and Graze-Out.

For each data set, <u>first start by combining data from all the sites into one set.</u>

A. Initial Exploration: (20% effort) – Use 100% data for this task.

1. Conduct exploratory data analysis to better understand the relation between weather, EVI, and flux attributes.
2. Perform structural equation modeling (SEM) and spearman correlation analysis (SCA) to examine the direct and indirect effects of main environmental factors and EVI on flux attributes.
3. Identify key attributes from 'A' that can be leveraged to build modeling for each of the flux attributes (ET, NEE, ER, GPP).

B. Predictive Modeling to assess ecosystem dynamics and its key drivers: (60% effort) – – Use 80% data for training the models.

4. For each flux attribute, build predictive regression models (ML and DL) using weather and EVI.
5. Then, explore if adding other flux attributes contributes to improved models' performance using the below proposed approach:
   a. For modeling ET, explore using NEE.
   b. For modeling NEE, explore using ET.
   c. For modeling ER, explore using NEE and/or ET.
6. Also, Identify the key important input variables that are driving the modeling.

The key point here is to be able to model carbon and water dynamics and understand their controlling factors.

<u>Note:</u> This will be four independent modeling simulations, followed by three more analyses exploring the inclusion of other complimentary fluxes.

C. Summarize: (20% effort): -- Use 20% of data for testing the models.

7. From all the exploratory and modeling analysis for the two datasets (Grain-Only and Graze-Out), summarize and present your findings with recommendations and strategies to improve crop management practices.

For model evaluation, use standard metrics like MAE, MSE, RMSE, MAPE, R^2.

<u>Grading:</u>

25%: Extended abstract; 50%: presentation; 20%: coding.

The winner for each challenge will provide further support for code transfer, support in manuscript preparation and will be credited with co-authorship.