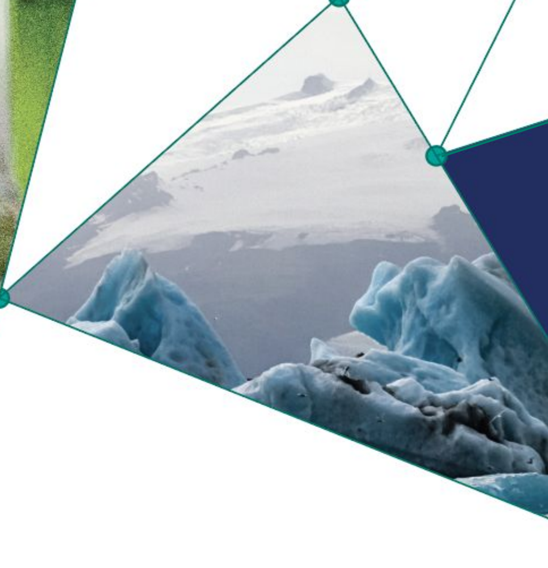
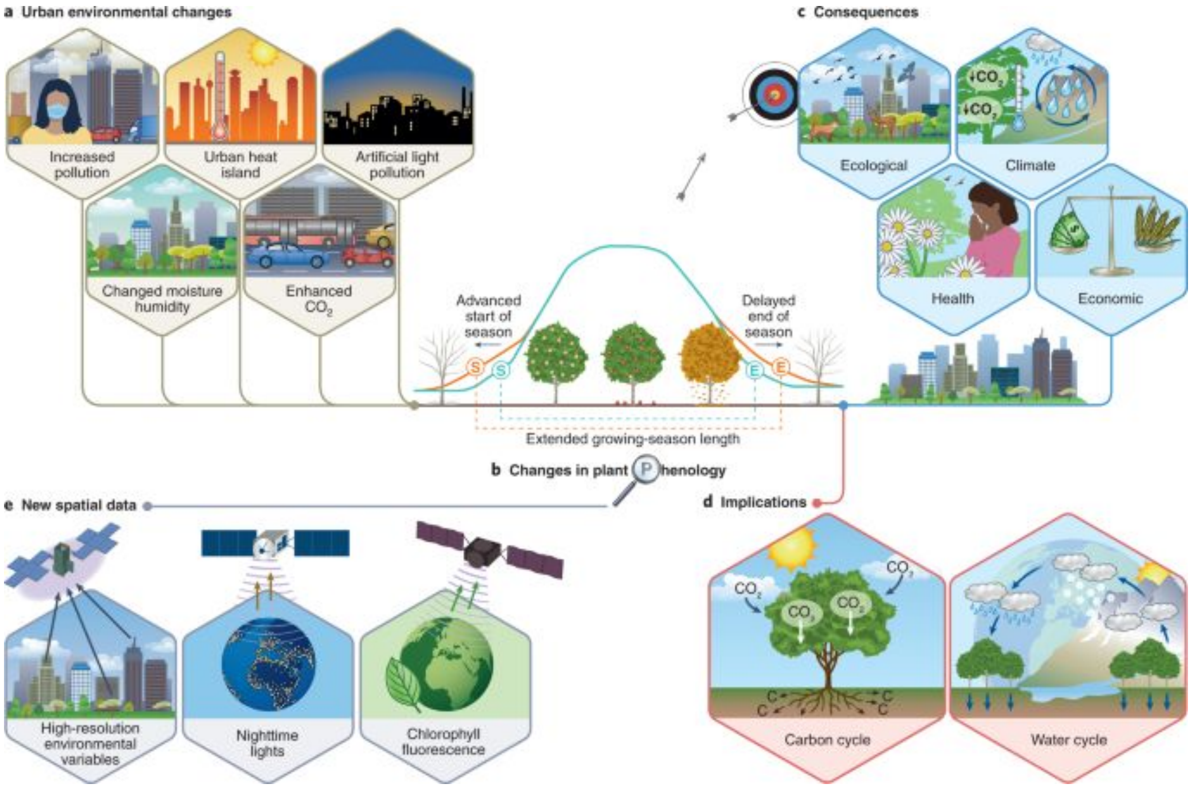


Towards Climate-Adaptive Phenology in CLM Using Machine Learning

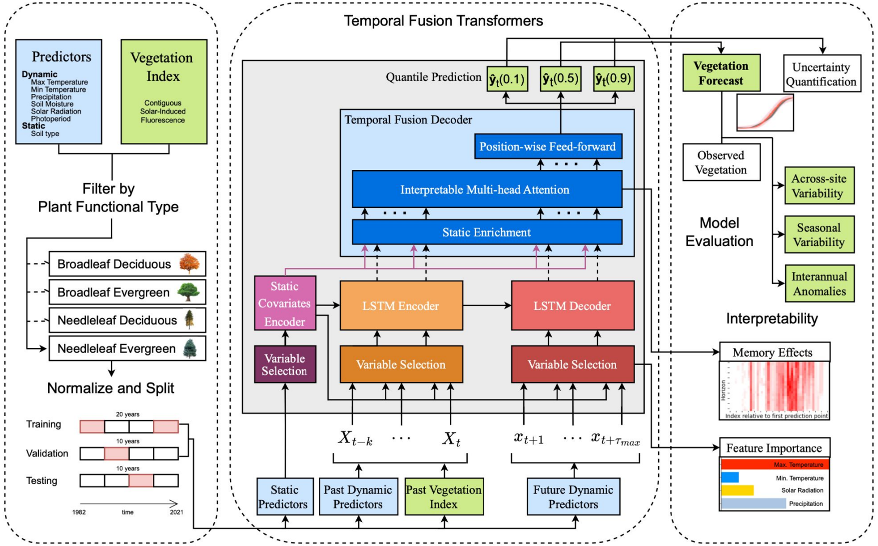
Aya Lahlou, Columbia University
Linnia Hawkins, Adrianna Foster, Dave Lawrence, Pierre Gentine
June 9th, 2025



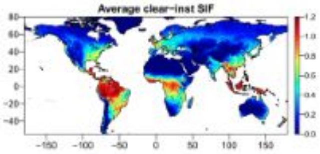
a Urban environmental changes
Increased pollution
Urban heat island
Artificial light pollution
Changed moisture humidity
Enhanced CO_2
c Consequences
Ecological
Climate
Health
Economic
Advanced start of season
Delayed end of season
Extended growing-season length
b Changes in plant Phenology
e New spatial data
High-resolution environmental variables
Nighttime lights
Chlorophyll fluorescence
d Implications
Carbon cycle
Water cycle

(Zhou, 2022)

Temporal Transformers for Phenology



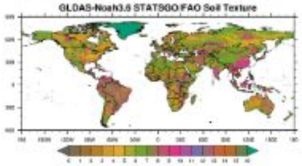
Predictand: LCSIF



- Long-term global spatially Contiguous Solar-Induced Fluorescence (Fang et al. 2022)
- 0.05 - 15 day resolution

Predictors:

- Minimum Temperature (Era5)
- Maximum Temperature (Era5)
- Accumulated Solar Radiation (Era5)
- Precipitation (Era5)
- Soil Moisture (Era5)
- Photoperiod
- Soil Type



Training set up:

- Resolution : 0.25 degrees - 1 day
- Historical context: 365 days
- Forecast Horizon: 30 days

Temporal Transformers Results

Figure 1: CSIF predicted (y axis) versus CSIF observed (x axis) with performance evaluated for all data points, for across-site variability, for seasonal variability, and for interannual anomalies. Seasonal variations (FSEAS) were derived by subtracting mean value per site from mean seasonal cycle. Anomalies were defined as deviations from the mean seasonal cycle.

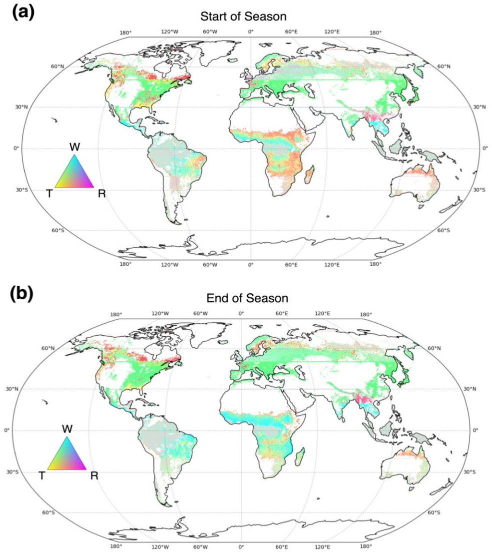
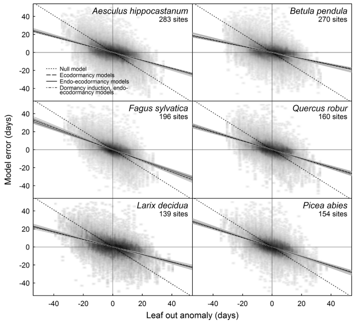


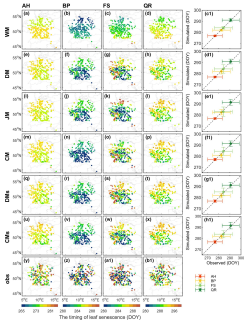
Figure 2: Interpretable TFT phenology predictions characterize EOS and SOS climate drivers and their spatial patterns

Poor Phenology Representation in the CLM and Beyond

Underestimated interannual variation of spring leaf phenology
Poor representation of inter-site variability of autumn leaf phenology



(Balser, 2016)



(Liu et al. 2020)

Bias in CLM5.0 Seasonal GPP vs FLUXCOM Observations

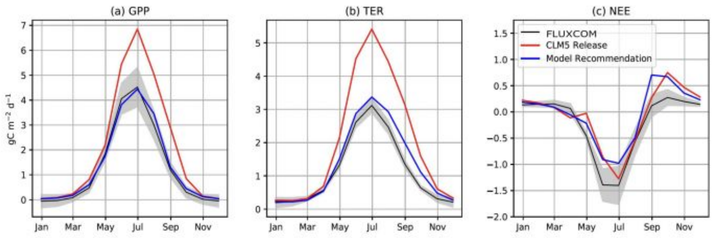


Figure 2. The annual cycle of **(a)** GPP, **(b)** TER, and **(c)** NEE from CLM5.0 and our model recommendation compared to FLUXCOM in the ABZ. Non-productive grid cells in the ABZ are removed from the average, meaning where LAI = 0, which is standard procedure in CLM analysis (Lawrence et al., 2019).

Best models having still have a 7 to 9 days prediction error

Empirical CLM Phenology: Limited sensitivity

Single on/off switches per year (no intra-seasonal variability)

Static parameters regardless of climate or PFT

LAI tracks C supply, not environmental forcing

Spring Onset

Initialization:

- Set Growing-Degree-Day sum: GDD_sum=0 \mathrm{GDD_sum} = 0GDD_sum=0

Soil Temperature GDD Trigger:

- If Tsoil3>273.15: GDD_sum += (Tsoil3-273.15)
- Compare to critical threshold: GDD_sum>**GDD_sum_crit** (function of Mean annual Temp), do onset

Onset Event:

- Trigger budburst in **CNPhenologyMod::cnonsetgrowth**
  - Fixed duration: 30 days
  - Transfers all stored C to leaves, linearly over 30 days
  - No mid-season drought response

Fall Senescence

Initialization:

- After summer solstice

Photoperiod Trigger

- If P<10.92P < 10.92P<10.92 h: begin senescence

Offset Routine:

- **CNPhenologyMod::cnoffsetlitterfall**
  - Fixed duration: 15 d
  - Leaf C → litter, linear rate
  - After 15 days: leaf C = 0, LAI = 0

Empirical CLM Phenology: Limited sensitivity

Why ML? Why not a better parametrizations?

- Tuning thresholds can help shift phenology but rarely improves year-to-year responsiveness

Phenological Realism

Climate-responsive growing-season duration

Mechanistic chilling and soil-moisture thresholds

Carry-over legacy state effects

Driver-responsive growth-rate functions

Site-specific phenological parameterization

Spring Onset

Initialization:

- Set Growing - Degree - Day sum:
  GDD_sum=0 \mathrm{GDD_sum} = 0GDD_sum=0

Soil Temperature GDD Trigger:

- If Tsoil3>273.15:
  GDD_sum += (Tsoil3−273.15)
- Compare to critical threshold:
  GDD_sum>**GDD_sum_crit** (function of Mean annual Temp), do onset

Onset Event:

- Trigger budburst in **CNPhenologyMod::cnonsetgrowth**
  - Fixed duration: 30 days
  - Transfers all stored C to leaves, linearly over 30 days
  - No mid - season drought response

Fall Senescence

Initialization:

- After summer solstice

Photoperiod Trigger

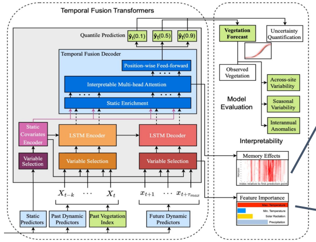
- If P<10.92P < 10.92P<10.92 h: begin senescence

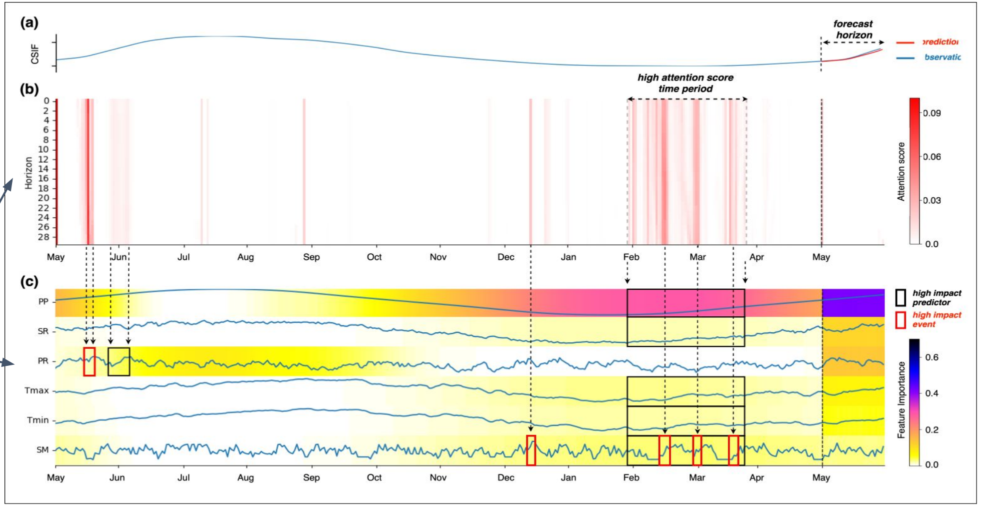
Offset Routine:

- **CNPhenologyMod::cnoffsetlitterfall**
  - Fixed duration: 15 d
  - Leaf C → litter, linear rate
  - After 15 days: leaf C = 0, LAI = 0

Temporal Transformers for Phenology

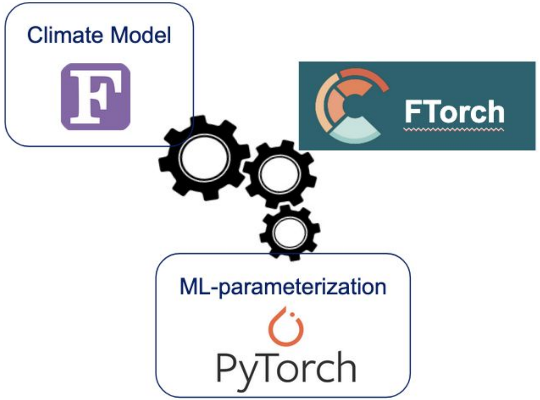
Attention → Long-Term multivariate Climate-Phenology effects





But ML models are in python, right?

- PyTorch models can be saved using existing Torch C++ interface
  - Compiled static PyTorch model in C++
- FTorch is an interface for C++ / Fortran built on LibTorch
  - FTorch library will be integrated in CESM3
- Using FTorch we can call a static PyTorch model
  - No python at runtime
- No online training (yet)



Proposed Method : Climate Sensitive SOS/EOS

May - July

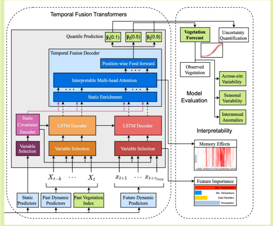
Phase 01

Inferring growth onset, maturation, senescence, dormancy from TFT predictions of LAI

60-day inputs

Tmin, Tmax, PP, Pr, SWC, SR

Trained TFT model weights

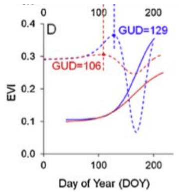


CLM

FTorch

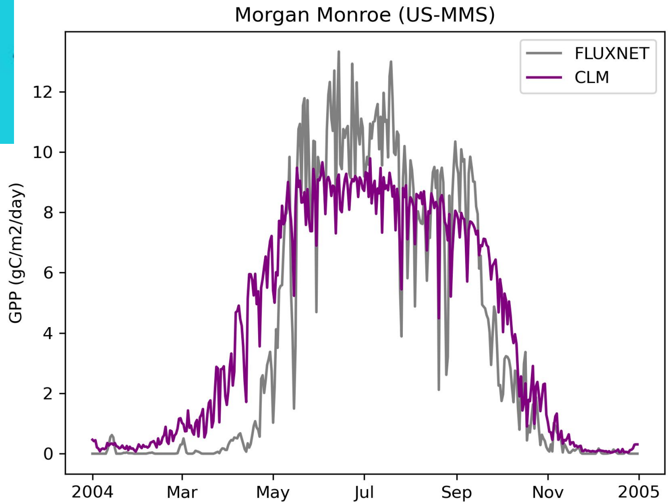
10-day LAI forecast signal

Inflection point detection: adaptive local iterative logistic fitting method



Preliminary evaluation: Morgan Monroe State Forest

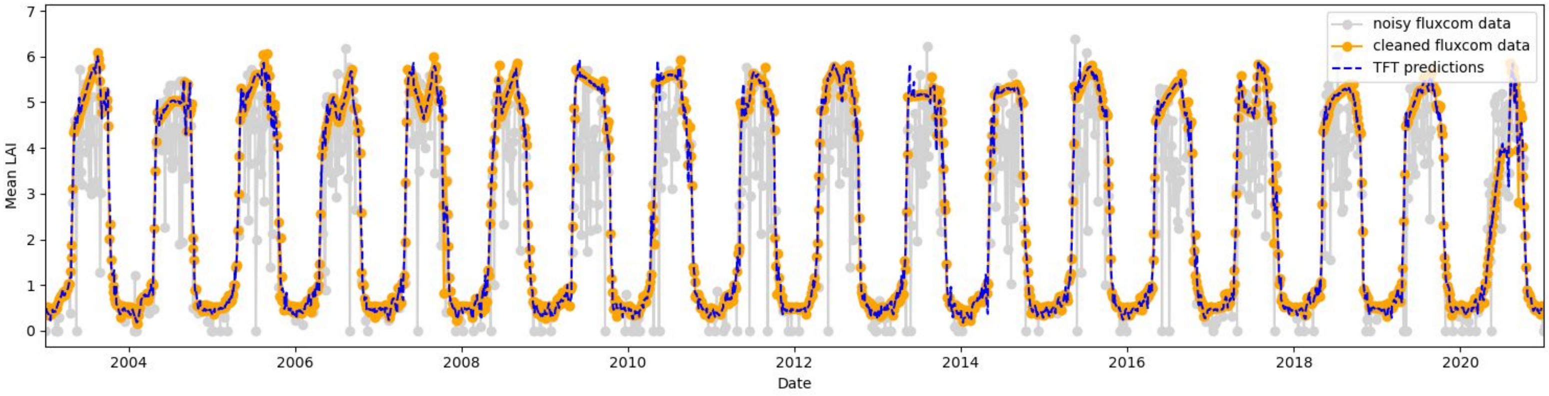
Deciduous Broadleaf forest, IN

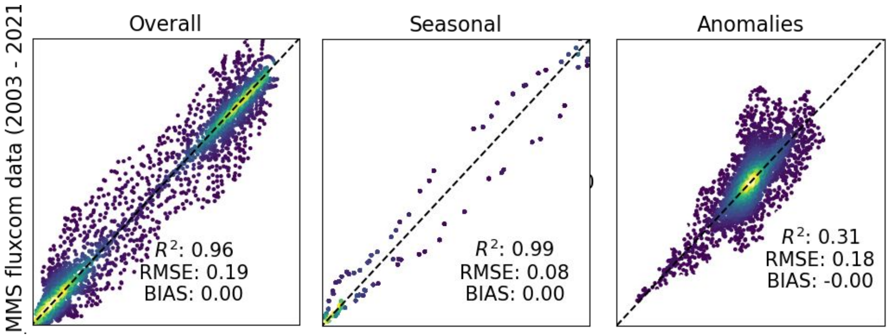


CLM:

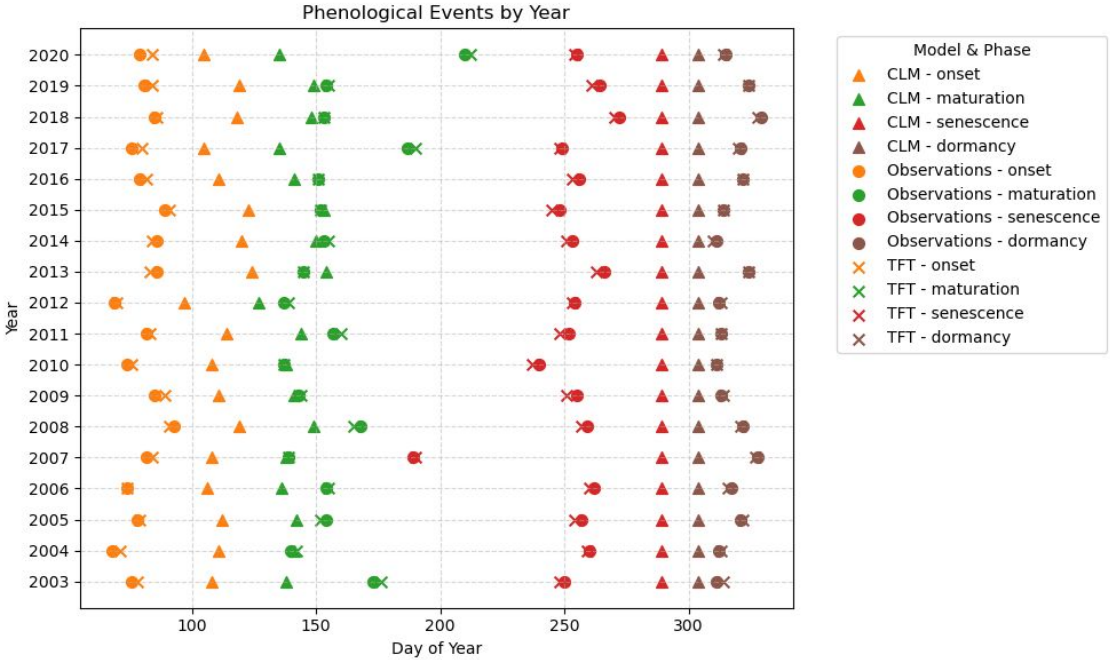
- Early onset
- Late senescence
- Consistent year-to-year

Preliminary results: US-MMS





Preliminary results: US-MMS



R^2 per Phase:

Observations vs TFT
onset: 0.840
maturation: 0.990
senescence: 0.978
dormancy: 0.969

Observations vs CLM
onset: -23.974
maturation: -0.835
senescence: -4.706
dormancy: -5.363

Proposed Method : Climate Sensitive Vegetation Dynamics

May - July

Phase 01

Inferring growth onset, maturation, senescence, dormancy from TFT predictions of LAI

Phase 02

Dynamic Carbon Allometry

Goal: Enable realistic climate sensitive seasonal LAI profiles

$$CF_{leaf\_stor,leaf\_xfer} = f_{stor,xfer} CS_{leaf\_stor} / \Delta t$$

$$CF_{froot\_stor,froot\_xfer} = f_{stor,xfer} CS_{froot\_stor} / \Delta t$$

$$CF_{livestem\_stor,livestem\_xfer} = f_{stor,xfer} CS_{livestem\_stor} / \Delta t$$

$$CF_{deadstem\_stor,deadstem\_xfer} = f_{stor,xfer} CS_{deadstem\_stor} / \Delta t$$

$$CF_{livecroot\_stor,livecroot\_xfer} = f_{stor,xfer} CS_{livecroot\_stor} / \Delta t$$

$$CF_{deadcroot\_stor,deadcroot\_xfer} = f_{stor,xfer} CS_{deadcroot\_stor} / \Delta t$$

$$CF_{gresp\_stor,gresp\_xfer} = f_{stor,xfer} CS_{gresp\_stor} / \Delta t$$

- ➢ CLM5’s leaf C allocation and LAI derive from static, hard-coded functions
- ➢ Lacks interannual or climate-driven flexibility (e.g. Season Length)
- ➢ Produces overly smooth, climate-insensitive intra-seasonal LAI dynamics

Proposed Method : Climate Sensitive Vegetation Dynamics

May - July

Phase 01

Inferring growth onset, maturation, senescence, dormancy from TFT predictions of LAI

Phase 02

Dynamic Carbon Allometry

Phase 03

Stress Deciduous Phenology, Evergreen Phenology, Grasslands, etc.

Key takeaways

Phenology is a key uncertainty in land modeling, affecting carbon, water, and energy cycles.

Current methods in CLM (and CMIP6 more broadly) lack the adaptability to capture changing climate responses, struggling with onset timing, season length, and carbon allocation variability

By replacing hard-coded rules with an adaptive ML phenology scheme, we can recover realistic, inter-annual and extreme-driven variability in leaf-out and senescence timing.

Real-time ML inference in CESM via FTorch integration

Thank you! Q&A ?