



**HABITERRE**



# SYMFONI - A "System-of-Systems" Solution to Quantify Carbon Outcome for Bioenergy Feedstock Production at the Field Level

*- an ARPA-E SMARTFARM Initiative*

**Kaiyu Guan, Ph.D.**

**Blue Waters Professor & Founding Director, Agroecosystem Sustainability Center  
University of Illinois at Urbana Champaign**

Email: [kaiyug@illinois.edu](mailto:kaiyug@illinois.edu), website: <http://faculty.nres.illinois.edu/~kaiyuguan/>

**Major Contributors:** Bin Peng, Evan DeLucia, Zhenong Jin, Wang Zhou, Sheng Wang, Jingyun Tang, Robert Grant, Chongya Jiang, Licheng Liu, Carl Bernacchi, DK Lee, Wendy Yang, Lisa Ainsworth, Ziqi Qin, etc



# ***SMARTFARM Vision:***



**Make it possible and profitable to optimize for yield and carbon intensity.**

**Technical Goal:  
Reliable, accurate,  
and cost-effective  
quantification of  
feedstock carbon  
intensity at the  
field level.**



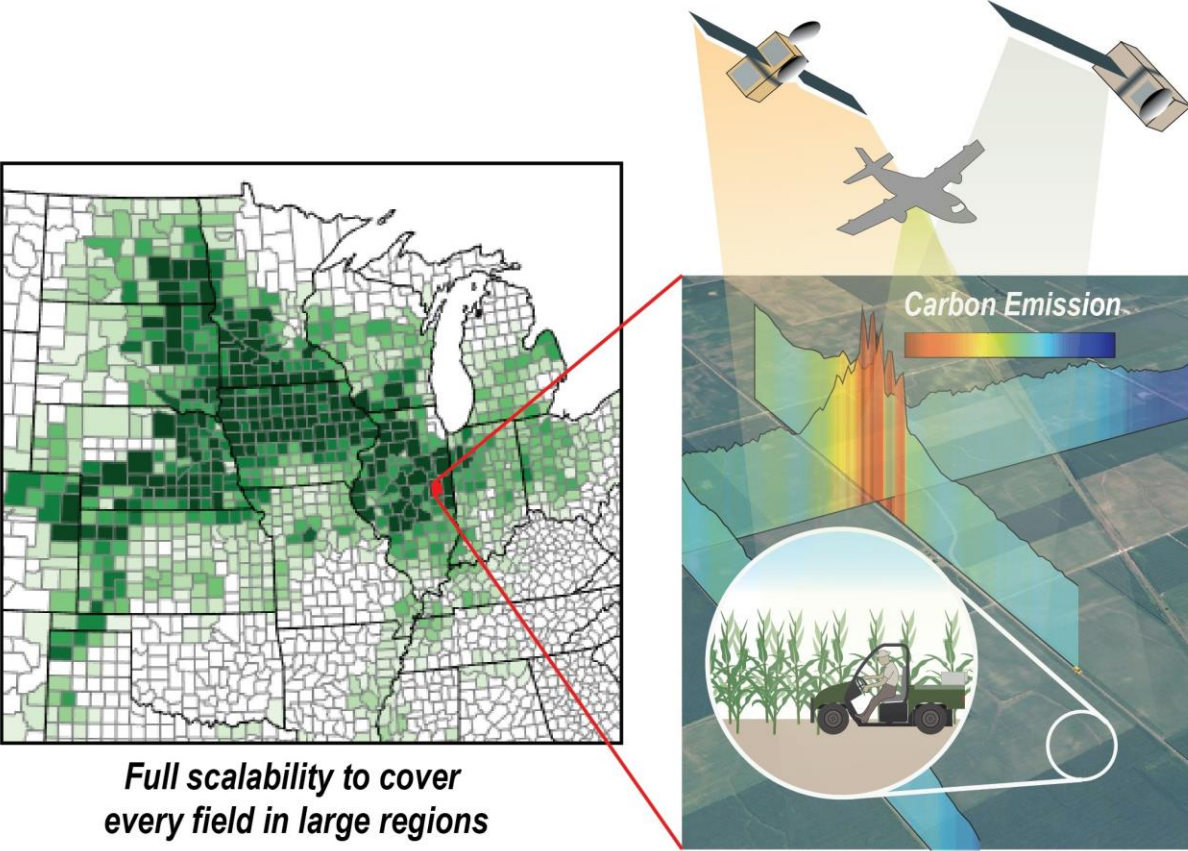
<b>David Babson</b>	<b>David Lee</b>	<b>Laura Demetron</b>
ARPA-E Program Director	ARPA-E SETA	ARPA-E SETA
David.Babson@hq.doe.gov	David.Lee2@hq.doe.gov	Laura.Demetron@hq.doe.gov

***Focus on emissions drivers (N<sub>2</sub>O) and net-negative strategies (soil carbon)***



# Our solution - "system of systems" approach:

Field accuracy + Scalability + Cost-effective



*Full scalability to cover every field in large regions*



## SYMFONI

A "system of systems" commercial solution of quantifying field-level carbon credit for farmland

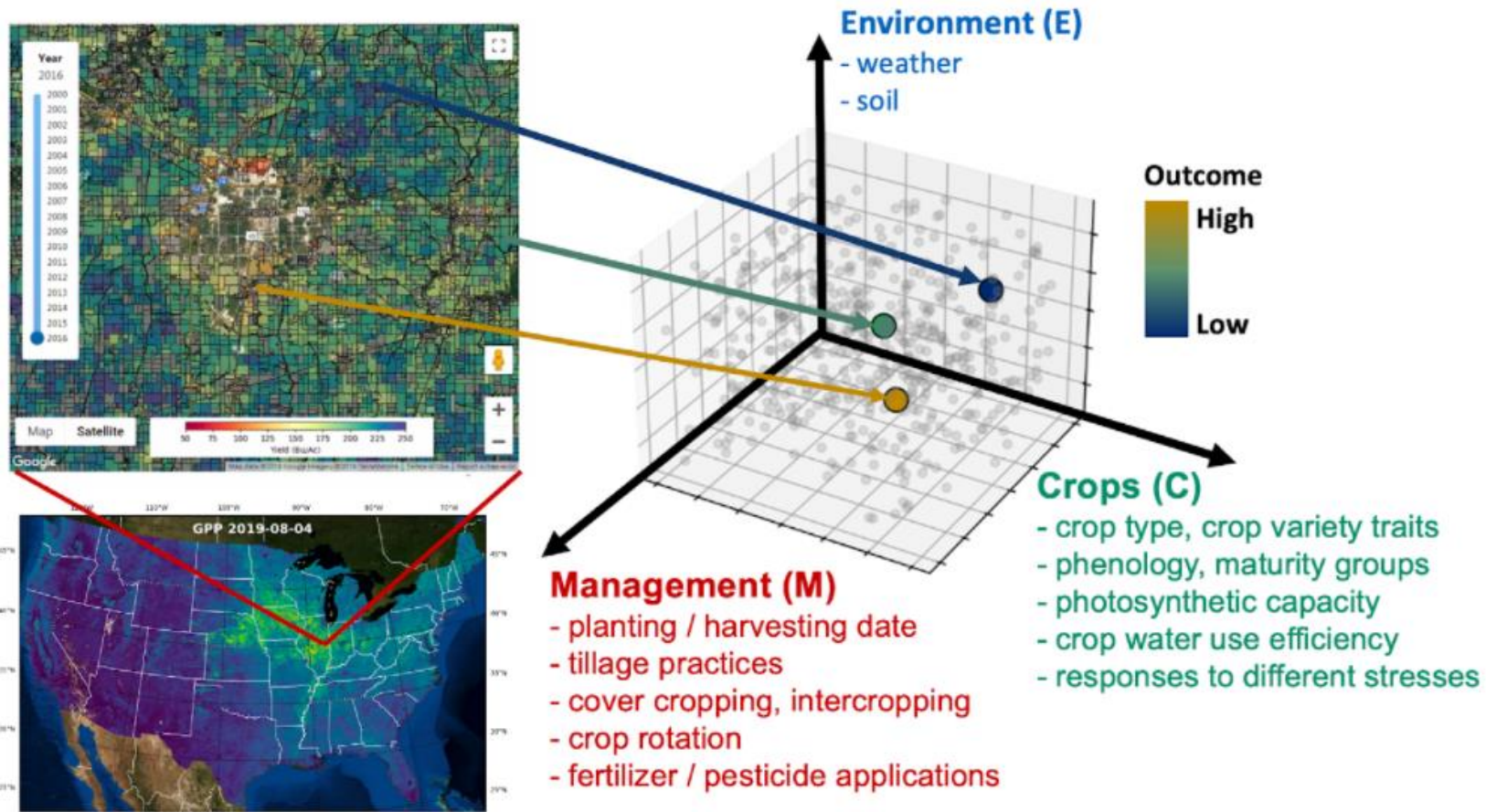
PI: Kaiyu Guan (University of Illinois)

 UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN	 UNIVERSITY OF MINNESOTA Driven to Discover™
 University at Buffalo	 BERKELEY LAB



# SYMFONI framework to quantify field-level carbon credit

(a) Agricultural Carbon Outcome = **Crops (C)** × **Management (M)** × **Environment (E)**





# Airborne-satellite integrated framework to quantify crop traits

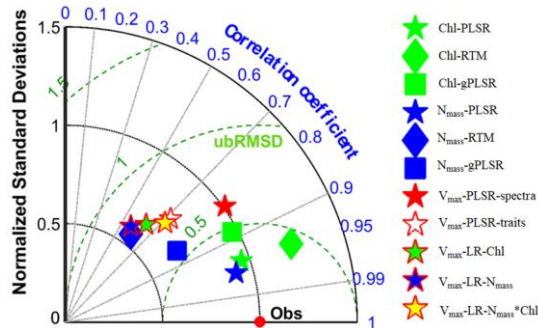
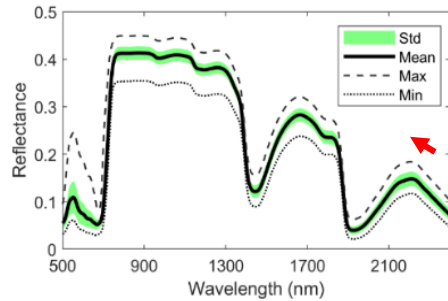
Hybrid approach to utilize **machine learning** and **radiative transfer models** to quantify crop nitrogen across scales



**Leaf scale**

Accuracy for corn nitrogen

$R^2 = 0.93$   
rRMSE = 6.83%

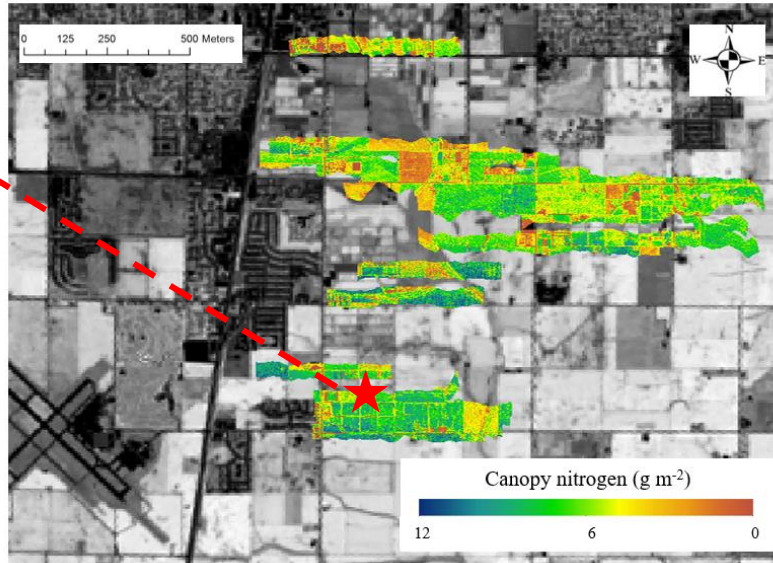


**Leaf reflectance**



**Airborne scale**

$R^2 = 0.85$   
rRMSE = 9.63%

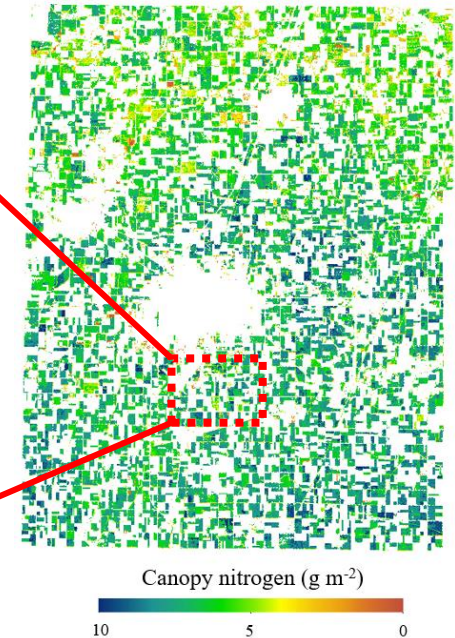


**UIUC South Farm fields**



**Satellite scale**

$R^2 = 0.72$   
rRMSE = 15.67%

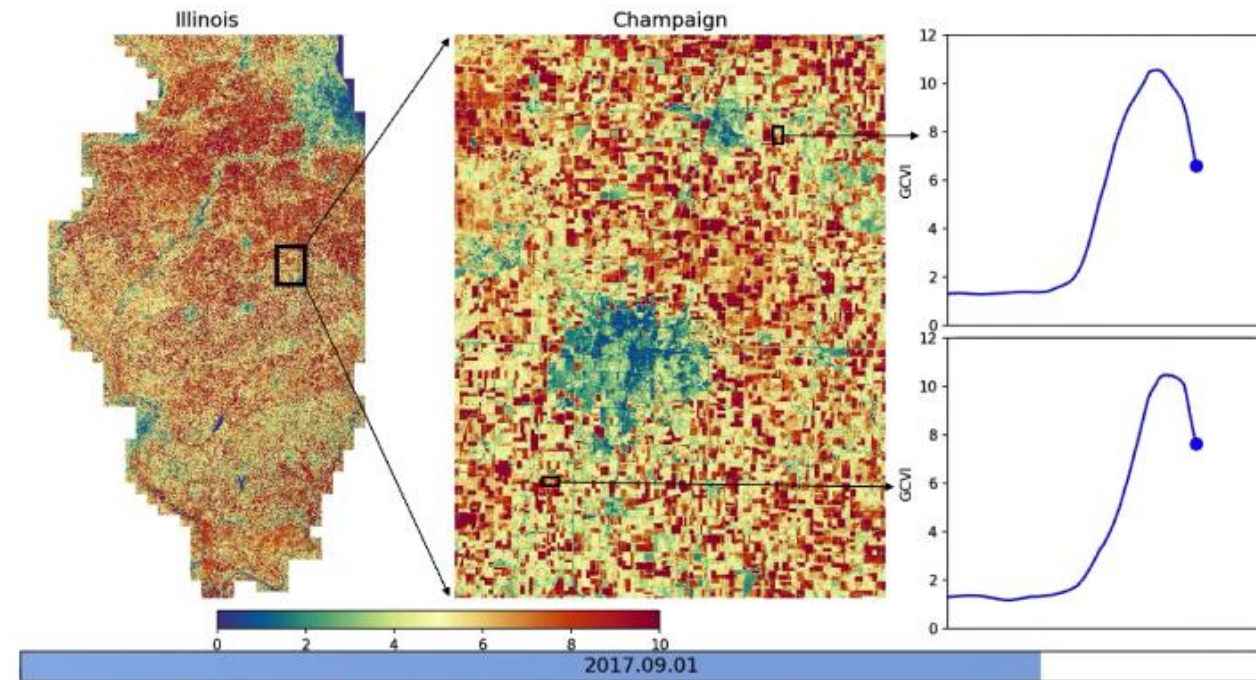
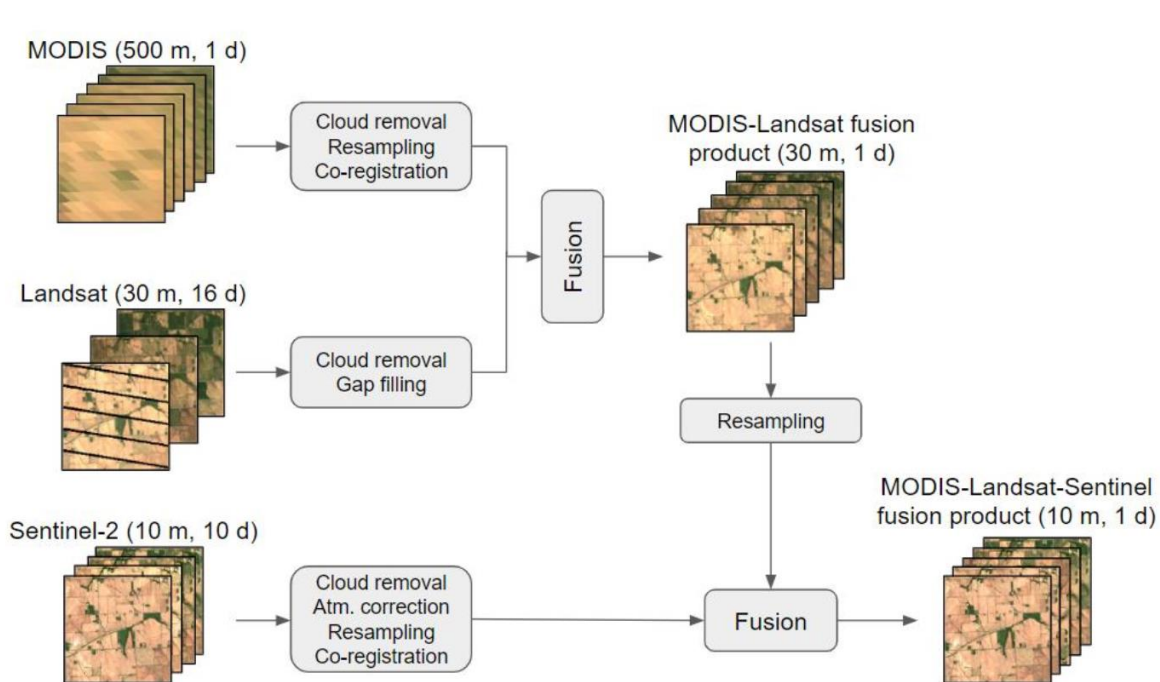


**Champaign county**

# Sensing & Monitoring:

## 2. Satellite fusion technology, to enable large-scale deployment

Scalable fusion algorithm combines various public satellite data to generate daily, 10-30 meter, gap-free/cloud-free images from 2000 to present.



**Luo, Y., Guan, K.\***, and Peng, J.\* (2018) "STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product", *Remote Sensing of Environment*.

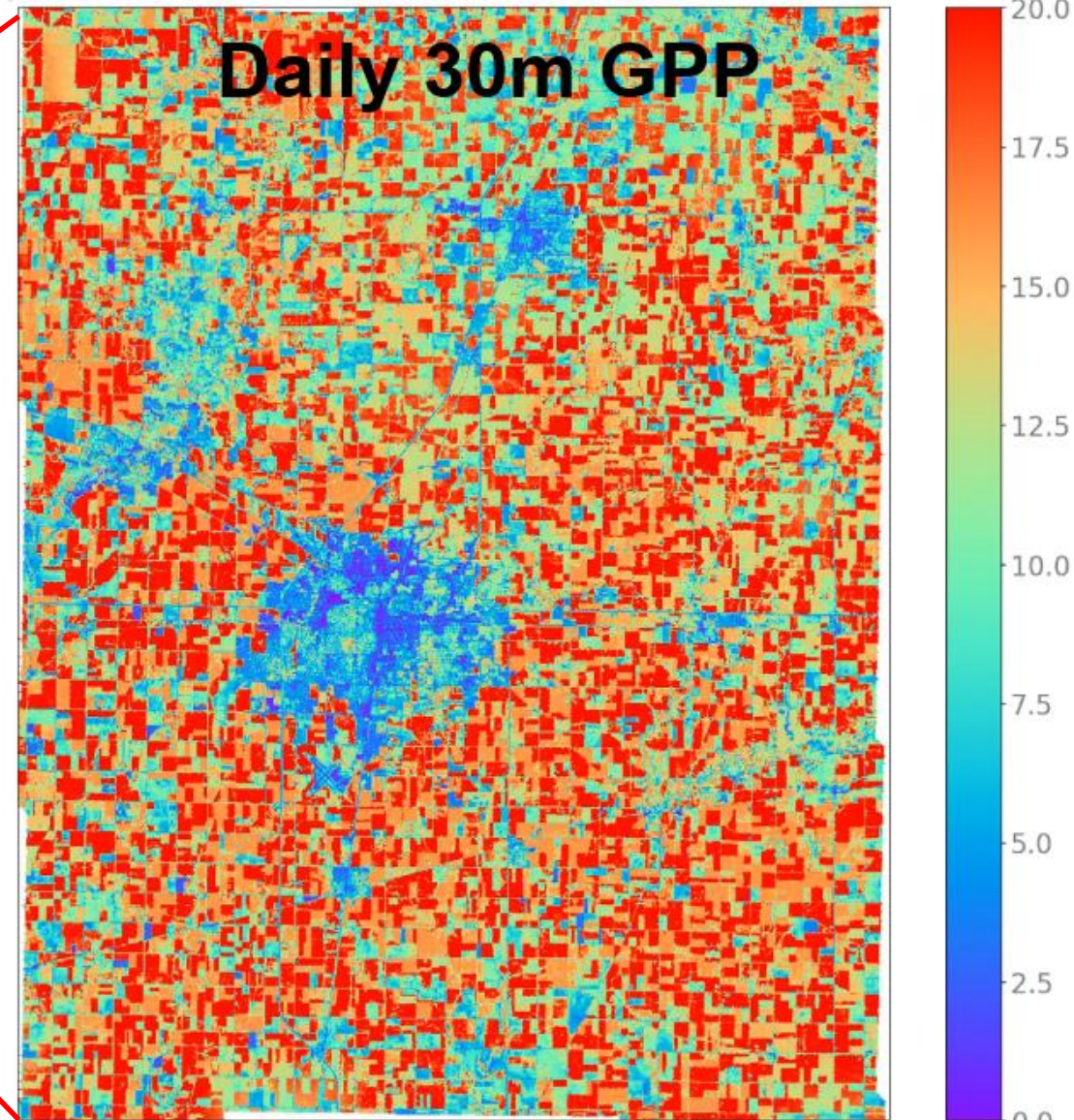
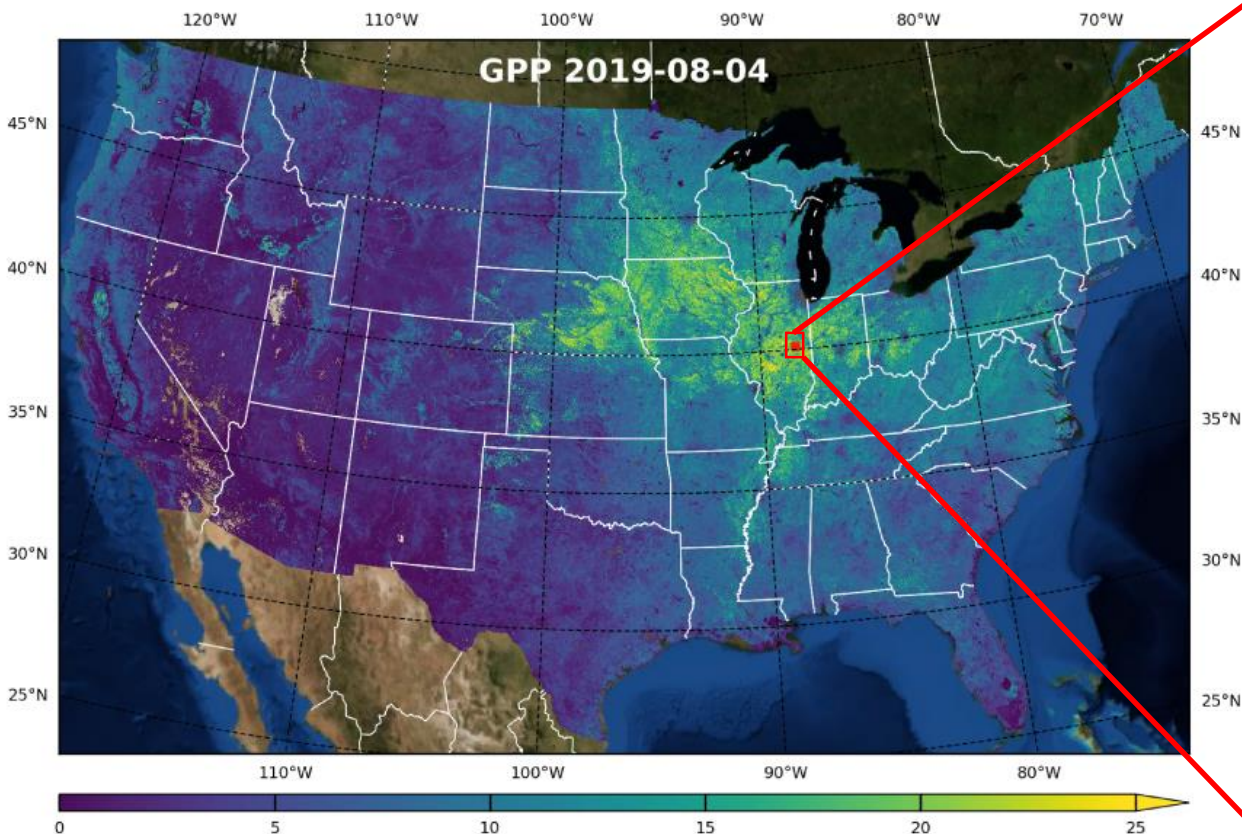
Check the demo video on YouTube: <https://youtu.be/IXVEVWIMQg4>

(PI: Guan, NASA NIP, 2016)



# Farmland Carbon Uptake (i.e. GPP, Photosynthesis)

10 meter resolution daily, real-time, covering every farmland in the United States



Jiang, C.\*, Guan, K.\*, Wu, G., Peng, B., and Wang, S. (2020) "A daily, 250 m, and real-time gross primary productivity product (2000–present) covering the Contiguous United States", Earth System Science Data.

Wu, G., Guan, K.\*, Jiang, C.\*, Peng, B., Kimm, H., Chen, M., Yang, X., Wang, S., Suyker, A.E., Bernacchi, C. and Moore, C.E. (2019) "[Radiance-based NIRv as a proxy for GPP of corn and soybean](#)", Environmental Research Letters.

Champaign, IL (2017)



You also heard of using models.

1. Quantifying the carbon outcome is complex enough that requires process-level understanding.

**Models need to have the necessary processes!**

2. The biggest problem of how models are used is:

**NO CONSTRAINTS to the model!**



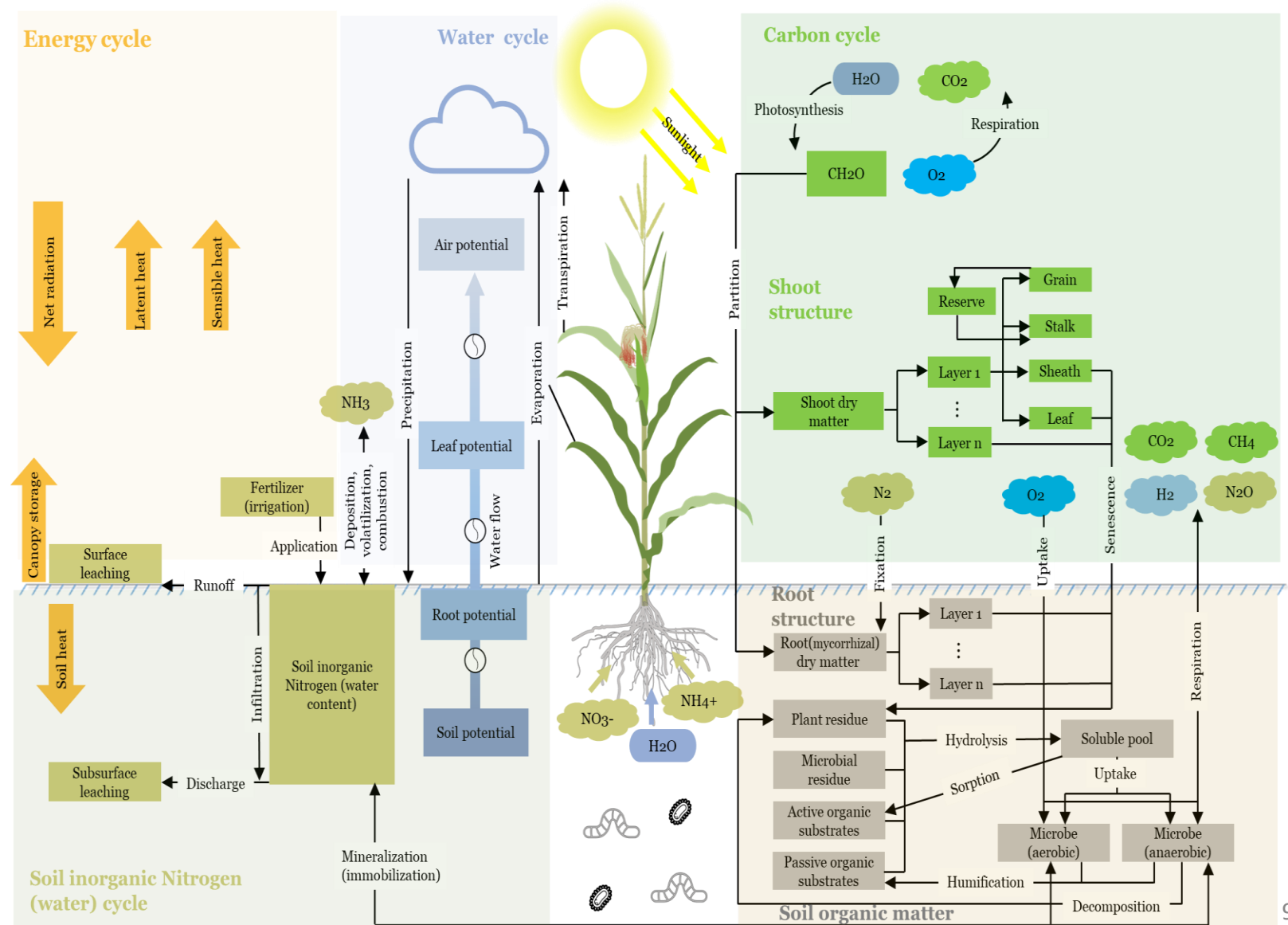


# Modeling & Quantification: Model-Data Fusion

## Ecosys model

*Ecosys* is an advanced mechanistic ecosystem model developed to simulate water, energy, carbon, and nutrient cycles simultaneously for various ecosystems at the hourly step with multiple soil and canopy layers. It can simulate the influence of different crop management practices (i.e. tillage, fertilizer, and irrigation) under different environmental conditions.

(<https://ecosys.ualberta.ca/>)





You also heard of using models.

1. Quantifying the carbon outcome is complex enough that requires process-level understanding.

**Models need to have the necessary processes!**

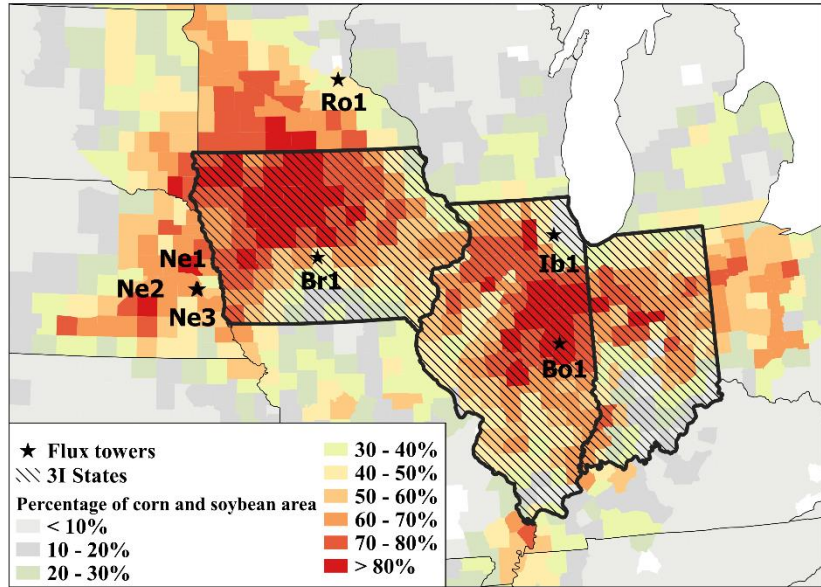
2. The biggest problem of how models are used is:

**NO CONSTRAINTS to the model!**





# Performance of Simulated Carbon Budget at Flux Tower Sites

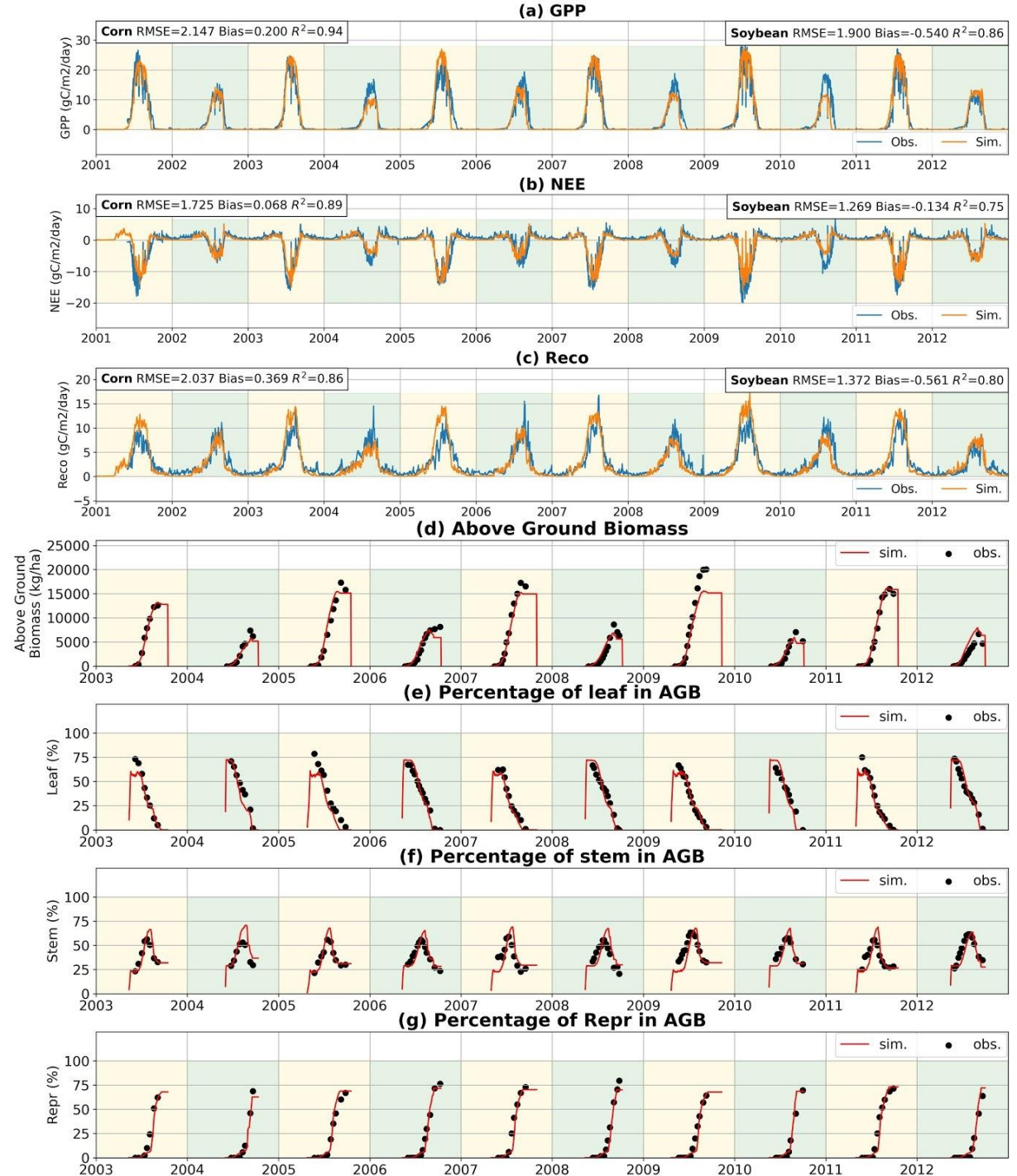


Locations of the seven flux towers and the three I states in the U.S. Midwest.

Comparing *ecosys* simulated GPP, NEE, Reco, and carbon allocation with site observations at Mead Ne3 site in Nebraska for both corn (light yellow shaded) and soybean (light blue shaded).

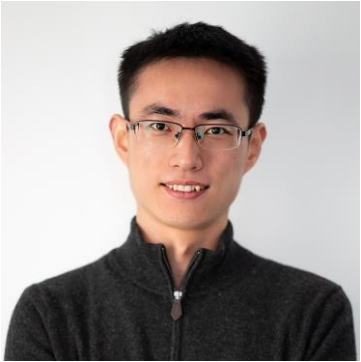
(Wang, Guan, et al. AFM, 2021)

(PI: Guan, NSF CAREER, 2020)



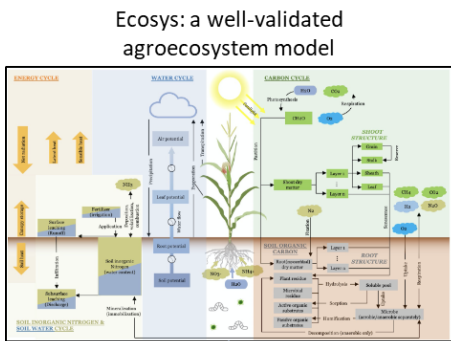


# AI-based data-model fusion technology (II) – Knowledge-Guided Machine Learning (KGML)



Zhenong Jin

Assistant Professor in Data Analytics for Precision Agriculture



Generate synthetic data

Synthetic data

Feature importance  
Causal relations

Design KGML structure

Initialized weights

Pre-train

Feature availability

Collect observed data



Source: Tim Griffis

Observed data

Refined KGML

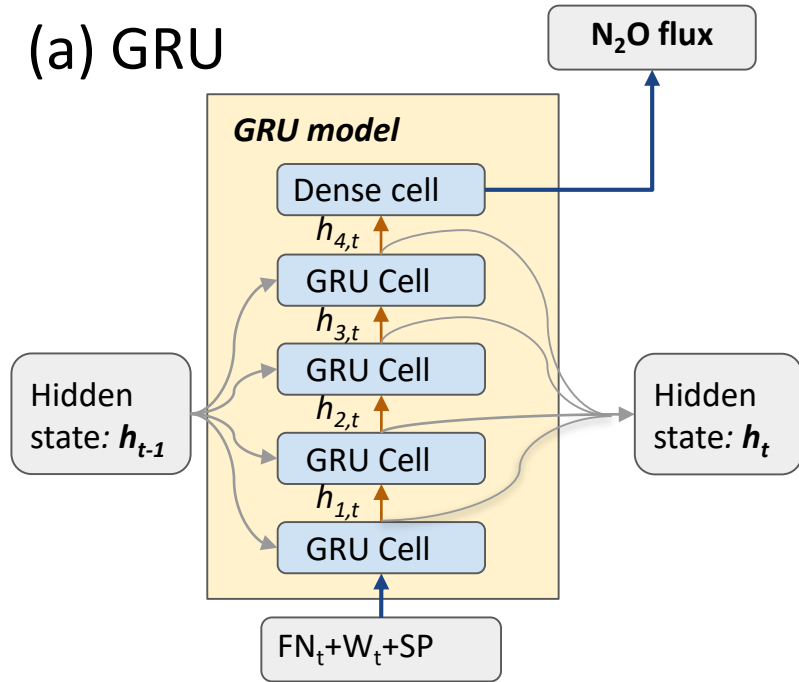
Re-train

Evaluate

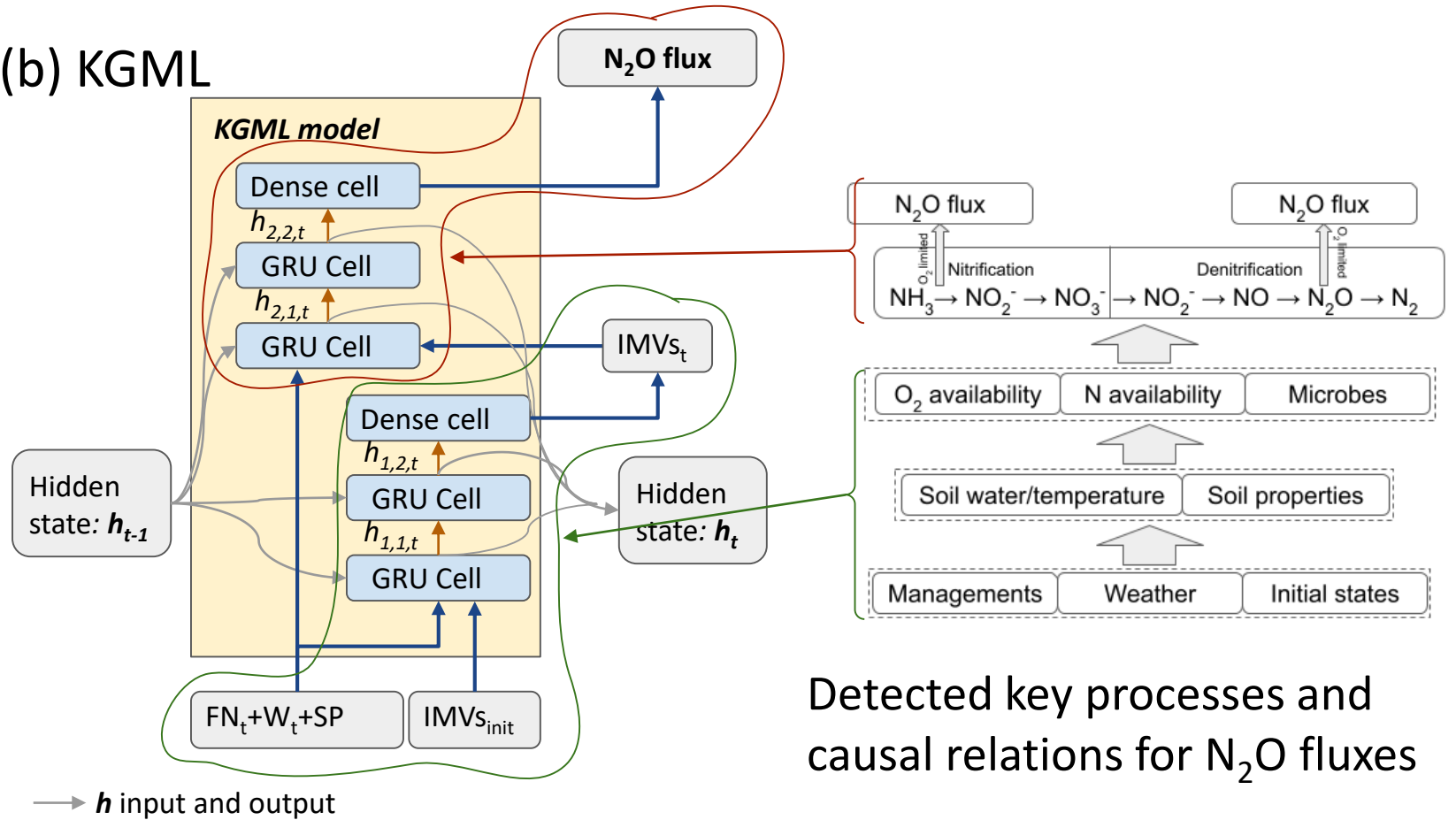
Liu L, ... Guan K, ..., Jin Z\* (2022) KGML-ag: A Modeling Framework of Knowledge-Guided Machine Learning to Simulate Agroecosystems: A Case Study of Estimating N2O Emission using Data from Mesocosm Experiments. **Geoscientific Model Development**

# Why KGML outperform others

(a) GRU



(b) KGML

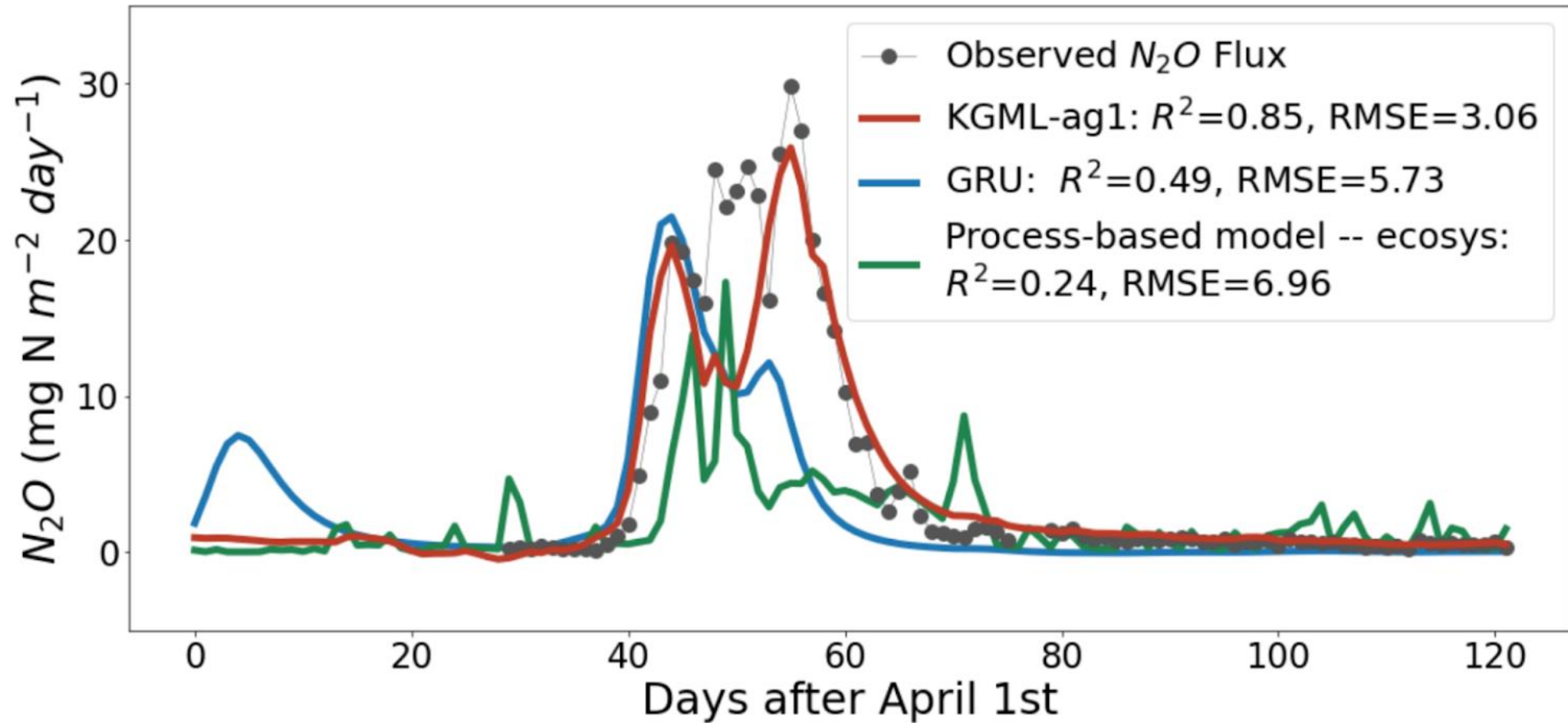


Detected key processes and causal relations for N<sub>2</sub>O fluxes

- GRU outperformed LSTM with its simpler structure in N<sub>2</sub>O simulation
- GRU model was used to do feature importance tests
- Physical guided initialization and architecture constraints were applied



## The KGML model outperformed the pure ML model



Liu L, ... Guan K, ..., Jin Z\* (2022) KGML-ag: A Modeling Framework of Knowledge-Guided Machine Learning to Simulate Agroecosystems: A Case Study of Estimating N<sub>2</sub>O Emission using Data from Mesocosm Experiments. *Geoscientific Model Development*

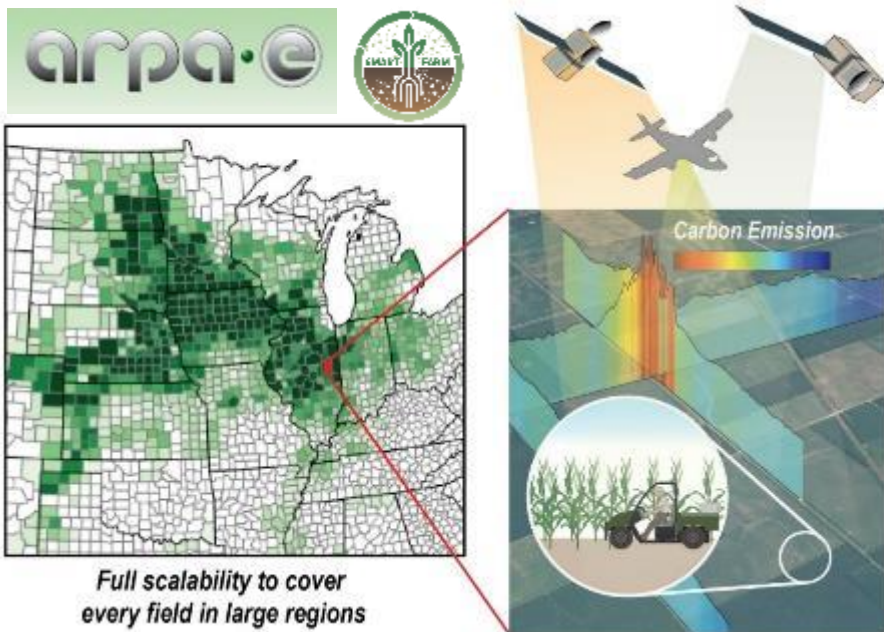


# HABITERRE

(previously “Aspiring Universe”)

- A UIUC-based startup that licensed the technology for commercialization

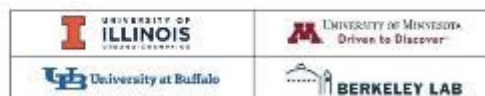
HabiTerre is the only Tech-to-Market (T2M) partner for the SYMFONI – the DOE ARPA-E SmartFarm grant for developing commercial-level carbon credit method.



## SYMFONI

A “system of systems” commercial solution of quantifying field-level carbon credit for farmland

PI: Kaiyu Guan (University of Illinois)



HabiTerre incorporates strong remote sensing capacity with advanced process-based modelling to achieve the lowest uncertainty, and utilizes AI to offer cost-effective scalable solutions.

Point of Contact: **Nick Reinke**, CEO of HabiTerre ([nick.reinke@habiterre.com](mailto:nick.reinke@habiterre.com))





**HABITERRE**



# SYMFONI - A "System-of-Systems" Solution to Quantify Carbon Outcome for Bioenergy Feedstock Production at the Field Level

*- an ARPA-E SMARTFARM Initiative*

**Kaiyu Guan, Ph.D.**

**Blue Waters Professor & Founding Director, Agroecosystem Sustainability Center University of Illinois at Urbana Champaign**

Email: [kaiyug@illinois.edu](mailto:kaiyug@illinois.edu), website: <http://faculty.nres.illinois.edu/~kaiyuguan/>

**Contributors:** Bin Peng, Evan DeLucia, Zhenong Jin, Wang Zhou, Sheng Wang, Jingyun Tang, Robert Grant, Chongya Jiang, Licheng Liu, Carl Bernacchi, DK Lee, Wendy Yang, Lisa Ainsworth, Ziqi Qin, etc.



**HABITERRE: Nick Reinke, CEO of HabiTerre ([nick.reinke@habiterre.com](mailto:nick.reinke@habiterre.com))**