Schedule for

5th Training course (primarily online with an option of in-person) on

New Advances in Land Carbon Cycle Modeling

(16-27 May 2022 with weekend on 21 and 22 off)

Note: All time in this schedule refers to the time in **Phoenix, Arizona, US** (UTC-7)

Overview: This training course will have 10 units. Units 1-5 are on the matrix approach to land carbon cycle modeling. Attendee will learn how to convert land carbon cycle models to matrix equations (units 1 and 2), add diagnostic variables in the models (unit 3), implement semi-analytic spin-up (SASU) to accelerate model initialization (unit 4), and conduct traceability analysis (unit 5). In addition, attendees will learn about general theories and essential concepts behind the matrix approach in units 1-5. Units 6-10 are on data assimilation, ecological forecasting, machine learning and deep learning. Attendees will learn essential concepts and the seven-step procedure of data assimilation (unit 6), conduct data assimilation with data from field studies and remote sensing (unit 7), perform ecological forecasting (unit 8), apply machine learning in carbon cycle research (unit 9), and integrate deep learning with data assimilation to improve predictions (unit 10).

Design of training units: We strive to design lectures to be understandable and practices to be doable by attendees with minimal background in modeling. Meanwhile, we offer options for attendees with advanced modeling skills to work on specific projects. We are particularly ready to help you apply the matrix approach to your own models or conduct data assimilation and machine learning with your own data. Meanwhile, nine invited talks will present you the latest results on relevant research subjects.

Materials available for training: In support to this training course, we have prepared pre-recorded lectures (videos) and a book, *Land Carbon Cycle Modeling: Matrix Approach, Data Assimilation, and Ecological Forecasting*. The book also has 10 units. Each unit has three chapters and one practice. We have prepared a pre-recorded lecture and a pre-recorded instruction for each chapter and each practice, respectively. Note that the content of units 7 and 8 in the book will be merged in unit 7 and the content of unit 9 in the book will be taught in unit 8 in 2022. We added machine learning in carbon cycle research as unit 9 in 2022. We have also prepared three additional chapters and corresponding pre-recorded preparatory lectures on linear algebra, Python programming, and the software, CarboTrain, developed for the practice sessions of the training course.

Methods: The 5th training course will primarily be held online. We also offer an option for in-person interactions if attendees plan to travel to Flagstaff, AZ, USA for two weeks. Both online and in-person attendees will take the course mostly in asynchrony in combination with synchronized (i.e., live virtual) meetings via Zoom. For each unit, attendees will attend one synchronized virtual meeting, read three book chapters (and supplementary materials, if any), listen to three pre-recorded lectures, take a quiz for each lecture, and do one practice according to pre-recorded instruction (most of the practices will be done with CarboTrain). Our instructors will go over your answers to quizzes and practices to ensure that you

understand the concepts and master the skills. In each synchronized virtual meeting, we will have one invited talk (starting from unit 2), assess your learning in the previous unit, give an overview on the current unit, and answer your questions.

We offer in-person interactions with attendees who travel to Flagstaff. Specifically, we will have in-person discussion session with those in-person attendees in each afternoon during the training course. Note that in-person attendees will study the materials in the same way as the online attendees.

Attendees who do not have much background in linear algebra and/or Python programming need to study the preparatory lectures as described in the two appendix chapters in the book before 16 May, 2022.

Appendix 1 Ye Chen: Matrix algebra in land carbon cycle modeling
Appendix 2 Xin Huang: Brief introduction on programming in Python

Important: All attendees need to install CarboTrain following Appendix 3 before the training course starts.

Appendix 3 Yuan Gao: CarboTrain user guide for Carbon cycle modeling Training course (**Note that R version needs to be updated to 4.1.0 or higher**).

Alternatively, you can run a docker container by using the following link. The CarboTrain has been configured and is ready to use as long as you can run the docker container by following the instructions in the page.

https://hub.docker.com/r/gaoyuan199325/carbotrain

All attendees need to attend the opening meeting on the first day and the virtual meeting for each unit via Zoom unless the time is not good for you (the Zoom meetings will be recorded and shared with attendees). In addition, we offer a virtual mixing via Zoom each day (attending this virtual mixing is optional).

Times: Attendees go through 10 units at your locations, one unit per day from 16-27 May 2022 (with weekend on 21 and 22 off). Attendees may need about 4 hours to finish one unit.

The opening meeting via Zoom (7:00-8:00am on 16 May)
Introduction by attendees and instructors by groups
Yiqi Luo: Overview and objectives of the training course
Yiqi Luo: Overview of unit 1

Unit 1: Fundamentals of carbon cycle modeling

Pre-recorded lectures (* indicates the lecture is essential for attendees to do practices)
L1.1 *Yiqi Luo: Theoretical foundation of the land carbon cycle and matrix approach
L1.2 Ben Smith: Introduction to carbon cycle modeling

L1.3 *Yuanyuan Huang: Flow diagrams and balance equations of land carbon cycle models

Pre-recorded instruction for practice
P1 Yuanyuan Huang: Developing carbon flow diagrams and balance equation

Synchronized mixing via Zoom (5:00pm on 16 May)

Unit 2: **Matrix representation of carbon balance**

Synchronized virtual meeting (7:00-8:00am on 17 May)
- Invited talk by Pep Canadell
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L2.1 *Yuanyuan Huang: Developing matrix representation of land carbon models
L2.2 Zheng Shi and Xingjie (Chris) Lu: Coupled carbon-nitrogen matrix models
L2.3 Carlos Sierra: Compartmental systems

Pre-recorded instruction for practice
P2 Yuanyuan Huang: Matrix representation of carbon balance equations and coding

Synchronized mixing via Zoom (5:00pm on 17 May)

Unit 3: **Carbon cycle diagnostics for uncertainty analysis**

Synchronized virtual meeting (7:00-8:00am on 18 May)
- Invited talk by Yang Song
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L3.1 *Yiqi Luo: Unified diagnostic system for uncertainty analysis
L3.2 Yuanyuan Huang: Sensitivity analysis with matrix equations: A case study with ORCHIDEE
L3.3 Enqing Hou: Matrix phosphorus model and data assimilation

Pre-recorded instruction for practice
P3 Xingjie (Chris) Lu: Diagnostic variables in matrix models

Synchronized mixing via Zoom (5:00pm on 18 May)

Unit 4: **Semi-analytic spin-up (SASU)**

Synchronized virtual meeting (7:00-8:00am on 19 May)
- Invited talk by Dave Lawrence, NCAR

- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L4.1 Ying Wang: Non-autonomous ODE system solver and stability analysis
L4.2 *Xingjie (Chris) Lu: Semi-Analytic Spin-Up (SASU) of coupled carbon-nitrogen cycle models
L4.3 Carlos Sierra: Time characteristics of compartmental systems

Pre-recorded instruction for practice
P4 Xingjie (Chris) Lu: Efficiency and convergence of semi-analytic spin-up (SASU) in TECO

Synchronized mixing via Zoom (5:00pm on 19 May)

Unit 5: **Traceability and benchmark analysis**

Synchronized virtual meeting (7:00-8:00am on 20 May)
- Invited talk by Kostiantyn Viatkin
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L5.1 *Jianyang Xia: Overview of traceability analysis
L5.2 Lifen Jiang: Applications of the transient traceability framework
L5.3 Forrest Hoffman: Benchmark analysis

Pre-recorded instruction for practice
P5 Jianyang Xia & Jian Zhou: Traceability analysis for evaluating terrestrial carbon cycle models

Synchronized mixing via Zoom (5:00pm on 20 May)

Unit 6: **Introduction to data assimilation**

Synchronized virtual meeting (7:00-8:00am on 23 May)
- Invited talk by Natasha MacBean
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L6.1 *Yiqi Luo: Data assimilation: Introduction, procedure, and applications
L6.2 *Feng Tao: Bayesian statistics and Markov chain Monte Carlo method in data assimilation
L6.3 Junyi Liang: Application of data assimilation to soil incubation data

Pre-recorded instruction for practice
P6 Xin Huang: The seven-step procedure for data assimilation

Synchronized mixing via Zoom (5:00pm on 23 May)

Unit 7: **Data assimilation with field measurements and satellite data**

Synchronized virtual meeting (7:00-8:00am on 24 May)

- Invited talk by Hyungsub Kim
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L7.1 Daniel Ricciuto: Model-data integration at the SPRUCE experiment
L7.2 *Shuang Ma: Application of data assimilation to a peatland methane study
L7.3 Matthew Williams: Global Data Assimilation Using Earth Observation – the CARDAMOM approach

Pre-recorded instruction for practice
P7 Shuang Ma: Assimilation at the SPRUCE site

Synchronized mixing via Zoom (5:00pm on 24 May)

Unit 8: **Ecological forecasting with EcoPAD**

Synchronized virtual meeting (7:00-8:00am on 25 May)

- Invited talk by Istem Fer
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L8.1 Yiqi Luo: Introduction to ecological forecasting
L8.2 *Shuang Ma: Ecological Platform for Assimilating Data (EcoPAD) for ecological forecasting
L8.3 Xin Huang: Cyberinfrastructure to facilitate ecological forecasting (no book chapter)

Pre-recorded instruction for practice
P8 Jiang Jiang: Ecological forecasting at the SPRUCE site

Synchronized mixing via Zoom (5:00pm on 25 May)

Unit 9: Applications of machine learning to carbon cycle research

Synchronized virtual meeting (7:00-8:00am on 26 May)

- Invited talk by Emma Stell
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L9.1 Yuanyuan Huang: Machine learning in carbon cycle research
L9.2 Yao Zhang: Long-short term memory network for Earth science
L9.3 Yu Jiang: Sensing, AI, and Robotics for Sustainable and Resilient Specialty Crops Production

Pre-recorded instruction for practice
P9 Feng Tao and Kostiantyn Viatkin: Applications of random forest to predict SOC content in Macedonia

Synchronized mixing via Zoom (5:00pm on 26 May)

Unit 10: **Process-based machine learning and data-driven modeling (PRODA)**

Synchronized virtual meeting (7:00-8:00am on 27 May)

- Invited talk by Peter Reich
- General assessment of learning for the last unit and overview on the current unit
- Q&A

Pre-recorded lectures
L10.1 Toby Hocking: Introduction to machine learning and neural networks
L10.2 *Feng Tao and Yiqi Luo: PROcess-guided deep learning and Data-driven modeling (PRODA)
L10.3 Umakant Mishra: Applications of machine learning to soil carbon research (no book chapter)

Pre-recorded instruction for practice
P10 Feng Tao: Deep learning to optimize parameterization of CLM5.

Synchronized mixing and closing via Zoom (5:00pm on 27 May)