3rd Training course (virtual) on

New Advances in Land Carbon Cycle Modeling

Tentative schedule

Overview: This virtual 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 and benchmark analysis (unit 5). In addition, attendees will learn about general theory and some essential concepts behind the matrix approach in units 1-5. Units 6-10 are on data assimilation, ecological forecasting, and machine learning. Attendees will learn essential concepts and seven-step procedure of data assimilation (unit 6), make an application to SPRUCE project (unit 7), evaluate value of different data sets (unit 8), conduct ecological forecasting (unit 9), and combine machine learning with data assimilation to improve prediction (unit 10).

Design of training units: We strive to design lectures to be understandable and practices 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 with your own data. Meanwhile, nine invited talks will present you the latest results on relevant research subjects.

Methods: The training will be mostly conducted in asynchrony in combination with synchronized (i.e., live virtual) meetings via Zoom. For each unit of the training, attendees will attend one synchronized virtual meeting, read one paper/syllabus, listen to pre-recorded lectures, take quizzes for each lecture, and do practices according to pre-recorded instruction. Our instructors will go over your answers to quizzes and practices to make sure you understand the concepts and master the skills. The synchronized virtual meeting will have one invited talk, assess your learning in the previous unit, give an overview on the next unit, and answer your questions.

All the attendees need to attend the opening meeting at the very beginning of the training course and the virtual meeting for each unit via Zoom unless the time is not good for you.

Times: Attendees go through 10 units of training at your locations, one unit per day from 20-31 July, 2020 with weekend on July 25 and 26 off. Attendees may need to spend 2-6 hours to finish one unit of training.

The opening meeting via Zoom (7:00am on 20 July, Mountain Standard Time, UTC -7)
Introduction by attendees and instructors

Yiqi Luo: Overview and objectives of the virtual training course
Xin Huang: Instruction to install Python and R
Yiqi Luo: Overview of training in unit 1

Unit 1: Basic concepts and structures of carbon cycle models, including carbon flow diagrams and balance equations

Suggested reading
R1.1 Luo YQ, Weng ES. (2011) Dynamic disequilibrium of the terrestrial carbon cycle under global change. *Trends in Ecology & Evolution,* **26**, 96-104.

Pre-recorded lectures (* indicates the lecture is essential for attendees to do practices)
L1.1 *Yiqi Luo: Theoretical foundation of carbon cycle in terrestrial ecosystems
L1.2 *Yuanyuan Huang: Carbon flow diagrams and balance equations of TECO, CLM, and ORCHIDEE
L1.3 *Xin Huang: Brief introduction on programming (Python and R)

Practice session (Yuanyuan Huang)
P1 Pre-recorded instruction for practice: Developing carbon flow diagrams and balance equations of TECO, CLM, and possibly other models

Unit 2: **Matrix representation of carbon balance equations and coding**

Suggested reading
R2.1 Huang YY, Lu XJ, Shi Z, Lawrence D, Koven C, Xia JY, Du ZG, Kluzek E, Luo YQ. (2018) Matrix approach to land carbon cycle modeling: A case study with Community Land Model. *Global Change Biology*, **24**, 1394-1404.

Synchronized virtual meeting (7:00am on 21 July, Mountain Standard Time, UTC -7)
- Invited talk by Philippe Ciais on near real time estimate of fossil CO_2 emissions during the COVID-19 and implications for the terrestrial carbon balance
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures
L2.1 *Yuanyuan Huang: Development of matrix models for TECO, CLM and/or ORCHIDEE
L2.2 *Ye Chen: Basic operation of matrix equations
L2.3 Carlos Sierra: Compartmental systems and general properties of matrix equations

Practice session (Yuanyuan Huang)
P2 Pre-recorded instruction for practice: Creating and coding matrix equations of TECO, CLM and/or possibly other models

Unit 3: **Diagnostics of carbon cycle with matrix models for uncertainty analysis**

Suggested reading

R3.1 Luo YQ, Shi Z, Lu X *et al.* (2017) Transient dynamics of terrestrial carbon storage: mathematical foundation and its applications. *Biogeosciences,* **14**, 145-161.

Synchronized virtual meeting (7:00am on 22 July, Mountain Standard Time, UTC -7)

- Invited talk by James Randerson
- General assessment of learning for the last unit and overview on the next 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 of ORCHIDEE

L3.3 Enqing Hou: Matrix phosphorus model and data assimilation

Practice session (Chris Lu)

P3 Pre-recorded instruction for practice: Adding diagnostic variables in matrix models for uncertainty analysis

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

Suggested reading

R4.1 Xia JY, Luo YQ, Wang YP, Weng ES, Hararuk O. (2012) A semi-analytical solution to accelerate spin-up of a coupled carbon and nitrogen land model to steady state. *Geoscientific Model Development,* **5**, 1259-1271.

Synchronized virtual meeting (7:00am on 23 July, Mountain Standard Time, UTC -7)

- Invited talk by Danica Lombardozzi on CLM5
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L4.1 *Ye Chen: Basic concepts and operation of ordinary differential equation

L4.2 *Chris Lu: Semi-Analytic Spin-Up (SASU) of coupled carbon-nitrogen CLM matrix model

L4.3 Carlos Sierra: Time characteristics of land carbon cycle

Practice session (Chris Lu)

P4 Pre-recorded instruction for practice: Adding a module to enable SASU in TECO

Unit 5: **Traceability and benchmark analysis**

Suggested reading

R5.1 Xia JY, Luo YQ, Wang YP, Hararuk O. (2013) Traceable components of terrestrial carbon storage capacity in biogeochemical models. *Global Change Biology, **19**, 2104-2116.*

R5.2 Collier, N., F. M. Hoffman, D. M. Lawrence, G. Keppel-Aleks, C. D. Koven, W. J. Riley, M. Mu, and J. T. Randerson (2018), The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation, *J. Adv. Model. Earth Syst*., 10(11):2731–2754, doi:10.1029/2018MS001354.

Synchronized virtual meeting (7:00am on 24 July, Mountain Standard Time, UTC -7)

- Invited talk by Stephen Sitch on TRENDY and model uncertainty
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L5.1 *Jianyang Xia: Overview of traceability analysis
L5.2 Enqing Hou: Traceability analysis with six matrix models at SPRUCE (Matrix MIP)
L5.3 *Forrest Hoffman: Overview of benchmark analysis

Practice session

P5.1 Pre-recorded instruction for practice (Jianyang Xia): TraceMe V1.0
P5.2 Pre-recorded instruction for practice (Nathan Collier): Benchmark analysis with ILAMB

Unit 6: **Introduction to data assimilation**

Suggested reading

R6.1 Xu, T., L. White, D. Hui, and Y. Luo. 2006. Probabilistic inversion of a terrestrial ecosystem model: Analysis of uncertainty in parameter estimation and model prediction. *Global Biogeochemical Cycles*, 20, GB2007, doi:10.1029/2005GB002468.

Synchronized virtual meeting (7:00am on 27 July, Mountain Standard Time, UTC -7)

- Invited talk by Matt Williams on data-model fusion
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L6.1 Yiqi Luo: Introduction to data assimilation
L6.2 *Yiqi Luo: Seven-step procedure for data assimilation as described by Xu et al. (2006)
L6.3 *Feng Tao: Introduction to Bayesian statistics and MCMC

Practice session (Xin Huang)

P6 Pre-recorded instruction for practice: Reproducing work in Xu et al. (2006) to learn the seven-step procedure for data assimilation

Unit 7: **Data assimilation with TECO at SPRUCE**

Suggested reading

R7.1 Shuang Ma, Jiang Jiang, Yuanyuan Huang, Zheng Shi, Rachel M. Wilson, Daniel Ricciuto, Stephen D. Sebestyen, Paul J. Hanson, Yiqi Luo. 2017. Data-Constrained Projections of Methane Fluxes in a Northern Minnesota Peatland in Response to Elevated CO2 and Warming. *Journal of Geophysical Research: Biogeosciences*, **122**: 2841-2861.

Synchronized virtual meeting (7:00am on 28 July, Mountain Standard Time, UTC -7)

- Invited talk by Sasha Hararuk on data assimilation
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L7.1 Daniel Ricciuto: Overview of SPRUCE and modeling activities
L7.2 *Shuang Ma: Applications of MCMC to DA at SPRUCE
L7.3 *Yuan Gao: EcoPAD training version and user instruction

Practice session (Shuang Ma)

P7 Pre-recorded instruction for practice: application of data assimilation with a more comprehensive TECO model to SPRUCE

Unit 8: **Values of data to constrain model and prediction**

Suggested reading

R8.1 Keenan TF, EA Davidson, JW Munger, AD Richardson. 2013. Rate my data: quantifying the value of ecological data for the development of models of the terrestrial carbon cycle. *Ecological Applications*, 23, 273-286

Synchronized virtual meeting (7:00am on 29 July, Mountain Standard Time, UTC -7)

- Invited talk by Belinda Medlyn on process-oriented data-model integration
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L8.1 *Enqing Hou: Information contents of different types of data sets to constrain parameters and predictions
L8.2 Lifen Jiang: Application of traceability analysis to Duke and Harvard Forests

Practice session (Enqing Hou)

P8.1 Pre-recorded instruction for practice: Evaluating values of different data sets in constraining model parameters and prediction

Unit 9: **Ecological forecasting with EcoPAD**

Suggested reading

R9.1 Jiang Jiang, Yuanyuan Huang, Shuang Ma, Mark Stacy, Zheng Shi, Daniel M. Ricciuto, Paul J. Hanson, Yiqi Luo. 2018. Forecasting responses of a northern peatland carbon cycle to elevated CO2 and a gradient of experimental warming. *Journal of Geophysical Research: Biogeosciences*, **123**: 1057-1071.

Synchronized virtual meeting (7:00am on 30 July, Mountain Standard Time, UTC -7)

- Invited talk by Quinn Thomas on ecological forecasting for lake clarity
- General assessment of learning for the last unit and overview on the next unit
- Q&A

Pre-recorded lectures

L9.1 Yiqi Luo: Introduction to ecological forecasting

L9.2 *Shuang Ma: General architecture of EcoPAD for ecological forecasting

Practice session (Jiang Jiang)

P9 Pre-recorded instruction for practice: Ecological forecasting at SPRUCE site

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

Suggested reading

R10.1 Feng Tao, Zhenghu Zhou, Yuanyuan Huang, Qianyu Li, Xingjie Lu, Shuang Ma, Xiaomeng Huang, Yishuang Liang, Gustaf Hugelius, Lifen Jiang, Russell Doughty, Zhehao Ren and Yiqi Luo. 2020. Deep Learning Optimizes Data-Driven Representation of Soil Organic Carbon in Earth System Model Over the Conterminous United States. *Frontiers in Big Data*, **3**:17. doi: 10.3389/fdata.2020.00017.

Synchronized virtual meeting (7:00am on 31 July, Mountain Standard Time, UTC -7)

- Invited talk by Markus Reichstein on machine learning and deep learning in geosciences
- General assessment of learning for the last unit and overview on the last unit
- Q&A

Pre-recorded lectures

L10.1 Toby Hocking: General concepts and application of machine learning

L10.2 Jarrett Barber: Available software/resources for machine learning

L10.3 *Feng Tao: Data assimilation with CLM5

Practice session (Feng Tao)

P10 Pre-recorded instruction for practice: Regional data assimilation with CLM5 matrix model and optimizing data-driven modeling with neural network in continental US

Instructors/lecturers

1. Nathan Collier, Oak Ridge National Laboratory, USA
2. Enqing Hou, Northern Arizona University, USA
3. Xin Huang, Northern Arizona University, USA

4. Yuanyuan Huang, CSIRO Oceans and Atmosphere, Australia
5. Jiang Jiang, Nanjing Forestry University, China
6. Cuijuan Liao, Tsinghua University, China
7. Chris Lu, Sun Yat-Sen University, China
8. Shuang Ma, Jet Propulsion Laboratory, USA
9. Zheng Shi, University of California Irvine, USA.
10. Feng Tao, Tsinghua University, China
11. Jianyang Xia, East China Normal University, China.

Lecturers

1. Jarret Barber, Northern Arizona University, USA
2. Ye Chen, Northern Arizona University, USA
3. Toby Hocking, Northern Arizona University, USA
4. Forrest Hoffman, Oak Ridge National laboratory, USA
5. Lifen Jiang, Northern Arizona University, USA
6. Yiqi Luo, Northern Arizona University, USA
7. Daniel Ricciuto, Oak Ridge National laboratory, USA
8. Carlos Sierra, Max Plank Institute of Biogeochemistry, Germany

Invited speakers for the synchronized virtual meetings

1. Philippe Ciais, Laboratoire des sciences du climat et de l'environnement, France
2. Sasha Hararuk, University of Central Florida, USA
3. Danica Lombardozzi, National Center for Atmospheric Research, USA
4. Belinda Medlyn, Western Sydney University, Australia
5. James Randerson, University of California Irvine, USA
6. Markus Reichstein, Max Plank Institute of Biogeochemistry, Germany
7. Stephen Stich, University of Exeter, UK
8. Quinn Thomas, Virginia Tech., USA
9. Matthew Williams, University of Edinburg, UK