Tackling Technical Challenges in Land Data Assimilation

June 14-16, 2021 · 9:00-12:00 EDT / 3:00 - 6:00 CEST / [Time Zone]
Virtual Workshop: Draft Program

Co-chairs[†]: Natasha MacBean (Indiana University), Jana Kolassa (NASA GMAO), Andy Fox (Joint Center for Satellite Data Assimilation), Tristan Quaife (University of Reading), Hannah Liddy (Columbia University/NASA GISS)

[†]Organized by the AIMES Land Data Assimilation Working Group

---

June 14

Applicability of data assimilation approaches across different land modeling groups

Meeting Link

8:50 AM EDT	Coffee/tea time to join the conversation early and test out your camera and microphone.
9:00 AM EDT	Welcome from the Co-Chairs: Introduction to the workshop context and goals
9:10 AM EDT	**Speaker 1: Patricia De Rosnay** (ECMWF) - Technical challenges of coupled land-atmosphere data assimilation for operational Numerical Weather Prediction and reanalyses
9:25 AM EDT	**Speaker 2: Eunjee Lee** (NASA GSFC) - Effect of land initialization on the skill of forecasting carbon fluxes on sub-seasonal to seasonal (S2S) time scales
9:40 AM EDT	**Speaker 3: Bertrand Bonan** (CNRM) - Monitoring land surface variables with LDAS-Monde: focus on assimilation approaches and applications to kilometric-scale spatial resolutions
9:55 AM EDT	*Break (5 minutes)*
10:00 AM EDT	**Speaker 4: Marko Scholze** (Lund University) - Experiences on terrestrial model parameter optimisation based from the Carbon Cycle Data Assimilation System using multiple observations

10:15 AM EDT	**Speaker 5: Breo Gomez** (UK Met Office) - Differences between atmospheric and land data assimilation and challenges for strong coupling
10:30 AM EDT	**Speaker 6: Sujay Kumar** (NASA GSFC) - Land hydrology data assimilation – Are we on the right track?
10:45 AM EDT	Introduction to Break Out Groups
10:50 AM EDT	*Break (10 minutes)*
11:00 AM EDT	Break Out Groups
11:45 AM EDT	Plenary Discussion/Report Backs
11:55 AM EDT	Co-chair wrap up

June 15

Emerging techniques
Meeting Link

8:50 AM EDT	Coffee/tea time to join the conversation early and test out your camera and microphone.
9:00 AM EDT	Welcome from the Co-Chairs: Introduction to Day 2
9:05 AM EDT	**Speaker 1: Jianzhi Dong** (USDA) - The added value of brightness temperature assimilation for global soil moisture estimation
9:20 AM EDT	**Speaker 2: Ewan Pinnington** (University of Reading) - Hybrid Data Assimilation Methods for Land Surface Modelling
9:35 AM EDT	**Speaker 3: Istem Fer (**Finnish Meteorological Institute) - Gaussian process emulators for efficient Bayesian calibration of process-based models
9:50 AM EDT	*Break (5 minutes)*
9:55 AM EDT	**Speaker 4: Joanne Waller** (UK Met Office) - Estimating the full observation error covariance matrix
10:10 AM EDT	**Speaker 5: Moha El Gharamti (**NCAR/UCAR) - Enhanced Streamflow Forecasting using Ensemble Data Assimilation
10:25 AM EDT	**Speaker 6: Anthony Bloom** (NASA JPL/Caltech) - Using an ever-growing Earth Observation record to infer and predict terrestrial C and H_2O dynamics
10:40 AM EDT	Introduction to Break Out Groups
10:45 AM EDT	*Break (10 minutes)*
10:55 AM EDT	Break Out Groups

11:40 AM EDT	Plenary Discussion/Report Backs
11:55 AM EDT	Co-chair wrap up

June 16

Challenges in dealing with observations
Meeting Link

8:50 AM EDT	Coffee/tea time to join the conversation early and test out your camera and microphone.
9:00 AM EDT	Welcome from the Co-Chairs: Introduction to Day 3
9:05 AM EDT	**Speaker 1: Nina Raoult** (LSCE) - Using the temporal dynamics of surface soil moisture to deal with biases when calibrating land surface models
9:20 AM EDT	**Speaker 2: Susan Steele-Dunne** (Delft University of Technology) - Towards constraining water and carbon cycle processes with radar data through assimilation
9:35 AM EDT	**Speaker 3: Jina Jeong** (Vrije Universiteit Amsterdam) - Using the International Tree-Ring Data Bank (ITRDB) records as century-long benchmarks for land-surface models
9:50 AM EDT	*Break (5 minutes)*
9:55 AM EDT	**Speaker 4: Ann Raiho** (Colorado State University) - Advances and challenges for using paleoecological data for state data assimilation within a forest gap model
10:10 AM EDT	**Speaker 5: Clara Draper** (NOAA) - Time scales in land data assimilation
10:25 AM EDT	**Speaker 6: Manuela Girotto** (UC Berkeley) - Technical challenges of assimilating observations with large spatiotemporal resolutions
10:40 AM EDT	Introduction to Break Out Groups
10:45 AM EDT	*Break (10 minutes)*
10:55 AM EDT	Break Out Groups
11:40 AM EDT	Plenary Discussion/Report Backs
11:55 AM EDT	Co-chair wrap up: Next steps

AIMES

analysis, integration, and modeling of the earth system

Tackling Technical Challenges in Land Data Assimilation

June 14-16, 2021 · 9:00-12:00 EDT / 3:00 - 6:00 CEST / Time Zone
Virtual Workshop: Abstract Booklet

Monday, June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
Invited Speakers: Patricia De Rosnay, ECMWF; Eunjee Lee, NASA/GSFC; Bertrand Bonan, Météo France; Marko Scholze, Lund University; Breo Gomez, UK Met Office; Sujay Kumar, NASA/GSFC

Tuesday, June 16, 2021: Challenges in dealing with observations
Invited Speakers: Jianzhi Dong, USDA; Ewan Pinnington, University of Reading; Istem Fer, Finnish Meteorological Institute; Joanne Waller, UK Met Office; Moha El Gharamti, NCAR/UCAR; Anthony Bloom, NASA/JPL

Wednesday, June 15, 2021: Emerging techniques
Invited Speakers: Nina Raoult, LSCE; Susan Steele-Dunn, TU Delft; Jina Jeong, Vrije Universiteit Amsterdam; Ann Raiho, Colorado State University; Clara Draper, NOAA ESRL; Manuela Girotto, UC Berkeley

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
9:10 AM EDT

Technical challenges of coupled land-atmosphere data assimilation for operational Numerical Weather Prediction and reanalyses

Patricia De Rosnay[1], Philip Browne[1], David Fairbairn[1], and Peter Weston[1]

[1]European Centre for Medium-Range Weather Forecasts

Abstract: This paper presents technical challenges of coupled land-atmosphere data assimilation for operational Numerical Weather Prediction (NWP) and reanalysis at the European Centre for Medium-Range Weather Forecasts (ECMWF).

ECMWF uses a weak land-atmosphere coupling approach, which relies on a coupled background forecast and separate land and atmosphere analyses. The ECMWF land data assimilation system uses a simplified Extended Kalman Fitter for the soil moisture analysis, and an Optimal Interpolation for screen level and snow data assimilation. In situ observations from the SYNOP network and satellite observations are assimilated in Near Real Time to provide land initial conditions to the forecast model. They are located at the land-atmosphere interface and include two-meter temperature and relative humidity, snow depth, and soil moisture related observations. I will describe the ECMWF land data assimilation system and briefly illustrate the impact of snow and soil moisture assimilation on the NWP performances.

I will present our activities related to modular infrastructure developments, to explore several coupling methods and to enable different coupling degrees to support consistent research and operations. I will introduce ongoing work to develop coupling through the observation operator for assimilation of satellite observations sensitive to the land-atmosphere interface conditions (e.g. snow). This requires consistent suite definition, file systems and observation interfaces for the land and atmosphere assimilation systems. Operational coupled assimilation requires having sustainable and near real time access to observations with consistent acquisition, pre-processing, archiving, and monitoring across the Earth system components. I will present these activities and discuss related practical and technical aspects.

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
9:25 AM EDT

Effect of Land Initialization on the Skill of Forecasting Carbon Fluxes on Sub-seasonal to Seasonal (S2S) Time Scales

Eunjee Lee[1,2], Randal. D. Koster[2], Lesley E. Ott[2], Joanna Joiner[4], Jana Kolassa[2,3], and Rolf Reichle[2]

[1]Goddard Earth Sciences Technology and Research, Universities Space Research Association, Columbia, MD, USA
[2]Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, MD, USA
[3]Science Systems and Applications, Inc., Lanham, MD, USA
[4]Atmospheric Chemistry and Dynamics Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD, USA

In this talk, we demonstrate an ability to forecast carbon fluxes accurately at multi-month leads in the Northern Hemisphere boreal region, an ability that appears to be linked in part to snowpack initialization in the forecast model. Using 20 years of forecasted meteorology from NASA GMAO's S2S ensemble forecast system (from forecasts initialized in December each year) to drive offline runs of the Catchment-CN model, we evaluate the degree to which we can forecast greening onset date and Gross Primary Production (GPP) relative to a fully independent observation dataset. We find that skillful forecasts of the greening onset date largely occur where we have skillful forecasts of snow cover removal date in western North America and Europe, and this snow cover removal date is naturally tied to the amount of snow at the start of the forecast. The ability to predict greening onset date in turn leads to an ability to predict annual (or more specifically, January-September) GPP, presumably because an earlier greening start date implies a longer time period for carbon uptake. We also, however, identify some regions for which GPP is accurately predicted without the snow mechanism; in these areas, it appears to be the initialization of the carbon/vegetation reservoirs in December that leads to the skill. Overall, this study demonstrates the significance of accurate land initialization for S2S carbon forecasts.

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
9:40 AM EDT

Monitoring land surface variables with LDAS-Monde: focus on assimilation approaches and applications to kilometric-scale spatial resolutions

Bertrand Bonan[1], Clément Albergel[2], Yongjun Zheng[1], Adrien Napoly[1], Alina L. Barbu[1], David Fairbairn[3], Simon Munier[1] and Jean-Christophe Calvet[1]

[1]CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
[2]European Space Agency Climate Office, ECSAT, Harwell Campus, Oxfordshire, Didcot OX11 0FD, UK
[3]European Centre for Medium-Range Weather Forecasts, Shinfield Road, Reading RG2 9AX, UK

Monitoring accurately the evolution of land surface variables (LSVs) such as soil moisture, biomass or leaf area index (LAI) is critical for various applications such as weather predictions, climate change or agricultural practices. Powerful instruments in that context are Land Data Assimilation Systems (LDASs) as they combine information from numerical simulations from Land Surface Models (LSMs) and satellite observations using data assimilation. One of them is LDAS-Monde, the offline LDAS developed by Météo-France's research centre (CNRM). It aims to monitor the evolution of LSVs at various scales, from regional to global. One specific feature of LDAS-Monde is to assimilate jointly satellite-derived observations of both surface soil moisture and LAI in order to update model estimations of LAI and soil moisture over the first metre of soil obtained with the multilayer and interactive vegetation ISBA LSM.

This presentation highlights two recent developments with LDAS-Monde. First, the impact of an Ensemble Kalman Filter (EnKF) and a Simplified Extended Kalman Filter (SEKF), that is routinely used in LDAS-Monde, on LSVs is assessed on an experiment over Europe and the Mediterranean basin at 0.25º spatial resolution. While the EnKF, newly applied in LDAS-Monde, exhibits on average a similar behaviour as the SEKF on LAI, it tends to have a positive impact on other LSVs such as surface soil moisture, evapotranspiration or gross primary production. The EnKF particularly improves simulations of river discharges obtained with the CTRIP river routing model coupled to ISBA. Secondly, the possibility to set-up LDAS-Monde to the context of kilometric-scale resolution is studied. To that end, satellite observations of LAI are assimilated into the ISBA LSM forced (offline mode) with Météo-France's small scale numerical weather prediction system AROME. It provides a system able to monitor LSVs at 2.5-km spatial resolution on a domain centred on France. The ability of LDAS-Monde to monitor the evolution of LSVs is demonstrated in the context of the severe drought that France suffered during the summer 2018 and 2020. LDAS-Monde's capacity to compensate the absence of irrigation in the ISBA LSM during summer by assimilating LAI is also shown over the Ebro valley. The presentation will focus on the technical challenges related to these experiments. In the future, merging those two recent developments is planned in order to create an ensemble-based LDAS operating at kilometre scale over France using ensemble atmospheric forecasts.

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
10:00 AM EDT

Experiences on terrestrial model parameter optimisation based from the Carbon Cycle Data Assimilation System using multiple observations

Marko Scholze[1]

[1]Lund University

In the context of climate change it is of paramount importance to understand CO2 sources and sinks and their spatio-temporal distribution. This information is needed to improve the projections of future trends in carbon sinks and sources, and thus the potential magnitude of climate change. However, there are large uncertainties in the quantification of the terrestrial carbon sinks arising mainly from uncertainties in the underlying models used for the quantification of these sinks. A major source for these model uncertainties are uncertainties in their parameterisations and parameter values. Reducing these uncertainties is critical for reducing the spread in simulations of the global carbon cycle, and hence in climate change projections. The Carbon Cycle Data Assimilation System (CCDAS) is designed to optimise model process parameters based on the assimilation of multiple data streams. Besides deriving an optimal set of parameters for the underlying process-based terrestrial biosphere model, here the BETHY model, a main feature of CCDAS is its capability of determining posterior parameter uncertainties consistent with the observational uncertainties of the assimilated data. These parameter uncertainties are then propagated onto the target quantities such as the net atmosphere surface exchange flux (NEP). In this presentation, I will report on the experiences gained in parameter optimisation from CCDAS studies assimilating multiple observations such as, e.g. CO2, FAPAR, soil moisture, vegetation optical depth.

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
10:15 AM EDT

Differences between atmospheric and land data assimilation and challenges for strong coupling

Breo Gomez[1]

[1]UK Met Office

Earth system modelling is part of the strategic plans of most numerical weather prediction (NWP) centres, including the Met Office. In recent years, significant effort has been put to implement a fully coupled atmosphere-ocean-land model in operations and the ambition is to extend this stronger interaction to data assimilation (DA) in the medium term. Integrating the DA components of two earth systems such as land and atmosphere presents a significant challenge as these typically use different scientific schemes and software packages. Even within the same earth system, there can be different approaches depending on the analysed variables, and such is the case for land surface DA. Soil variables (i.e. soil moisture) are typically analysed using Kalman Filter based algorithms and the analyses are usually produced independently for each soil column. Innovations are sometimes estimated in model quantities (i.e. soil moisture and snow products) and on other occasions in observation quantities (i.e. screen temperature values). Snow analyses follow a different approach to soil. Snow fields are commonly estimated using optimal interpolation algorithms which imply a global minimisation for the 2D field and snow innovations are typically calculated in observation space. On the contrary, atmospheric DA typically uses variational algorithms which process all the variables at the same time, over the full 3-D domain and over a time window (4-Dimensional). Innovations are computed in observation space and they ingest orders of magnitude more observations than land. All of these differences make reconciling the land and atmosphere systems a challenging task. Implementing a stronger coupling necessarily implies that land DA and atmosphere DA should eventually converge to a unified methodology and code base. In this talk, I will discuss the differences between atmospheric and land DA and outline some options to progress into a stronger coupling between the two.

June 14, 2021: Applicability of data assimilation approaches across different land modeling communities
10:30 AM EDT

Land hydrology data assimilation – Are we on the right track?

Sujay Kumar[1]

[1]NASA Goddard Space Flight Center

Physical modeling approaches are important tools for developing spatially and temporally continuous estimates of the terrestrial water cycle components. In addition to the limitations imposed by uncertainties in boundary conditions and model parameters, representing the heterogeneity and impacts of human management across different spatial scales are significant challenges in terrestrial hydrological modeling. The availability of remote sensing measurements offers the opportunity to mitigate some of these limitations, which are typically incorporated within physical models through data assimilation. Though assimilation studies have demonstrated the beneficial impact of remote sensing measurements both for improving the representation of catchment scale processes, there are significant challenges related to assimilation strategies, limitations in model formulations, and observational data processes that limit the potential utility of the remote sensing measurements. In this presentation, results from recent studies that highlight these challenges will be discussed. The information utilization within data assimilation approaches will be contrasted with data driven techniques based on machine learning. The use of nonparametric metrics founded in information-entropy methods that are more efficient for characterizing the information utilization efficiency will be discussed. The presentation argues for the need for hybrid approaches that integrate both physical and data-driven modeling to improve the representation of hydrologic processes across

June 15, 2021: Emerging techniques
9:05 AM EDT

The added value of brightness temperature assimilation for global soil moisture estimation

Jianzhi Dong[1]

[1]United States Department of Agriculture

Assimilation of microwave brightness temperature (TB) has been frequently used for improving global soil moisture estimates. However, the performance of data assimilation is sensitive to multitude factors, e.g., the reliability of the pre-defined observation and model error. As a result, assimilating brightness temperature does not necessarily improve soil moisture accuracy and for some cases, it may even degrade the baseline model run (open-loop, or OL). Therefore, the added value of TB assimilation is necessary for diagnosing and improving the land data assimilation systems. However, previous analyses have been confined to a few high-intensity networks and have limited global representativeness. Here, we developed statistical tools to evaluate the added value of TB data assimilation without any access to “ground-truth”. Based on the statistical methods, we evaluated the added value of TB assimilation for global soil moisture estimation in the SMAP L4 system. Results show that TB assimilation benefits soil moisture mostly in data-sparse regions, where OL is expected to have higher uncertainty levels due to inaccurate precipitation inputs, e.g., Africa and Central Australia. As for data-rich regions (e.g., Europe and US), data assimilation has marginal and even negative impacts on soil moisture estimates – suggesting SMAP L4 system should increase model weights to avoid soil moisture estimates being contaminated by remote sensing errors over these regions. We further showed that high-intensity soil moisture networks are generally located in data-rich regions. As such, traditional ground-based evaluations tend to underestimate the added value of TB assimilation at the global scale.

June 15, 2021: Emerging techniques
9:20 AM EDT

Hybrid Data Assimilation Methods for Land Surface Modelling

Ewan Pinnington[1]

[1]University of Reading

Land surface models often suffer from uncertain parameterisations and structural uncertainties. Variational data assimilation techniques can be optimal for the problem of parameter estimation with land surface models due to their ability to use a time series of observed data in a single minimisation and avoid the retrieval of unphysical time-varying parameters. Traditional variational data assimilation techniques such as Four-Dimensional Variational Data Assimilation require access to the model derivative and adjoint, these can be prohibitively costly to compute and maintain. In this talk we present a hybrid data assimilation technique that allows for the approximation of such variational techniques by using an ensemble of model runs. We will show examples of this technique applied to the JULES land surface model, assimilating NASA SMAP satellite observations, to optimize model soil parameters and improve hydrological predictions over the UK and Africa.

June 15, 2021: Emerging techniques
9:35 AM EDT

Gaussian process emulators for efficient Bayesian calibration of process-based models

Istem Fer[1], Michael Dietze[2]

[1]Finnish Meteorological Institute, Helsinki, FI
[2]Boston University, Boston, MA, USA

Bayesian calibration allows informing land surface models (LSMs) with data from multiple sources and scales, iteratively updating analyses as new data become available, propagating uncertainty into model predictions, and dealing with complex systems. While the primary aim of the calibration is constraining uncertainties in the model parameters, associated analyses help identify missing processes, feedback mechanisms or state variables.

The traditional Bayesian calibration algorithms, however, fail to leverage high-performance computing environments that are optimized for parallel computation and advances in computing power that are increasingly being made in terms of number of processors rather than CPU speed. This is more than an inconvenience where most LSMs are simply too slow to be plugged into these algorithms that require thousands to millions of sequential model evaluations.

To overcome this challenge we established an emulator-based Bayesian calibration framework where the emulator, that is orders of magnitude faster than the original computer simulator, is used in place of the full model and passed to Bayesian calibration algorithm. In this approach time limiting steps of running the full model are reduced and parallelized.

We use the Gaussian process(GP) model as our statistical emulator where GP always passes exactly through the design points, and allows for the estimation of uncertainties associated with interpolation in between design points. Key features of this approach involve emulating the error surface instead of model outputs, proposing and refining training points strategically, and modifying the calibration algorithm to accommodate for the uncertainty in GP.

The gains in terms of computation time using the emulator-based calibration are shown to be substantial with opportunities to explore more complex statistical models at the hierarchical level. We generalized and implemented the emulator-based Bayesian calibration and multi-site hierarchical Bayesian calibration work flows as part of an ecological informatics toolbox, PEcAn, where we make use of distributed architecture that facilitates community collaboration. We also discuss current limitations of the approach as well as potential solutions and more advanced applications that are under progress.

June 15, 2021: Emerging techniques
9:55 AM EDT

Estimating the full observation error covariance matrix

Joanne Waller[1]

[1]UK Met Office

For computational speed and simplicity the observation error covariance matrices used in geophysical data assimilation systems are typically assumed to be diagonal. However, recent research has shown that a variety of different observation types do have correlated errors; hence the diagonal error covariance matrix approximation is poor. Furthermore, using full observation error covariance matrices in the data assimilation scheme can inject small scale information into the analysis, reduce analysis errors and improve forecast skill. To attain these benefits, it is necessary to have an accurate estimate of the full observation error covariance matrix. In this presentation I will introduce some of the methods that can be used to estimate the observation error statistics. In particular I will discuss the advantages and deficiencies of a popular diagnostic that makes use of observation-minus-background and observation-minus-analysis statistics

June 15, 2021: Emerging techniques
10:10 AM EDT

Enhanced Streamflow Forecasting using Ensemble Data Assimilation

Moha El Gharamti[1]

[1]NCAR/UCAR

Predicting major floods during extreme rainfall events remain an important challenge. Rapid changes in flows over short time-scales combined with multiple sources of model error makes it difficult to accurately simulate intense floods. This work presents a general data assimilation framework that aims to improve flood predictions in channel routing models. The ensemble framework features: (i) along-the-stream (ATS), topology-based, localization and (ii) spatially and temporally varying adaptive inflation. Simulations show the importance of ATS localization compared to traditional approaches in mitigating not only sampling errors but also physically incorrect correlations between ungauged basins. Adaptive inflation on the other hand is utilized as a vigorous bias correction scheme in areas where the model's prediction skill is poor due to various biases and forcing errors. For a study case, hurricane Florence flooding in the Carolinas in 2018 is considered where hourly streamflow data – from USGS gauges – are assimilated. A regional configuration of WRF-Hydro (NOAA's National Water Model) is used to model the streamflow. The Data Assimilation Research Testbed (DART) is interfaced with WRF-Hydro to produce ensemble streamflow forecasts and analyses.

June 15, 2021: Emerging techniques
10:25 AM EDT

Using an ever-growing Earth Observation record to infer and predict terrestrial C and H_2O dynamics

Anthony Bloom[1]

[1]NASA Jet Propulsion Laboratory/Caltech

Understanding ecosystem-scale carbon (C) and water (H2O) exchanges is a challenging task, largely due to uncertainties in states, processes, C-H2O interactions, and their dynamic responses to climate variability and disturbance. In particular, both a process-based mechanistic understanding of C & H2O fluxes and a dynamical consideration of legacy effects on the terrestrial C and H2O states are needed to understand and predict the net land C sink in the coming decades. While land models explicitly represent these terms, considerable structural and parametric biases hinder efforts to accurately resolve C and H2O dynamics. The ever-growing Earth Observation record provides a unique opportunity for using observation to inform structural and parametric model configurations, and reduce associated uncertainties. Using the Bayesian CARbon DAta-MOdel fraMework (CARDAMOM) approach—constrained by an array of in-situ and satellite-based terrestrial ecosystem observations (including measurements of leaf area, biomass, solar-induced fluorescence, groundwater storage, satellite-informed estimates of CO2 and CO surface fluxes, and eddy covariance datasets)—we demonstrate the capacity of Earth Observations to jointly provide key insights into processes regulating the evolution of terrestrial ecosystem C fluxes on seasonal-to-decadal timescales. We present some of the key methodological steps—including spatially-explicit parametric inference, ecologically-based parameter regularization, joint state-parameter estimation, and explicit non-linear parameter optimization—which we argue are critical for accurately characterizing uncertainty in predictions of the land C sink in the coming decades.

June 16, 2021: Challenges in dealing with observations
9:05 AM EDT

Using the temporal dynamics of surface soil moisture to deal with biases when calibrating land surface models

Nina Raoult[1]

[1] LSCE

One of the biggest challenges we face when using surface soil moisture (SSM) data to evaluate and calibrate land surface models are the large biases that exist between modelled and observed values. These biases are further exacerbated when dealing with merged remote sensing products, such as the ESA CCI SM product, where the retrievals have been rescaled to match the climatology of an intermediate model. Methods to deal with these biases include preprocessing the data (e.g., cdf-matching) and assessing model-data fit through specialised metrics (e.g., unbiased root-mean squared error). Here we consider the novel approach of using only temporal dynamics of SSM to calibrate a model. Specifically, we consider 'drydowns' - the SSM temporal dynamics following a significant rainfall event - characterized by an exponential decay time scale $\boldsymbol{\tau}$. In this talk we also will discuss the important first step of selecting parameters through a sensitivity analysis. Using both the sensitivity analysis to calibrate only the most sensitive parameters and $\boldsymbol{\tau}$ approach in our calibration, we find that the relative drydowns of SSM can be well calibrated using observation-based $\boldsymbol{\tau}$ estimates. Furthermore, we investigate the overall model skill improvement using eddy-covariance measurements and discuss next steps.

June 16, 2021: Challenges in dealing with observations
9:20 AM EDT

Towards constraining water and carbon cycle processes with radar data through assimilation

Susan Steele-Dunne[1]

[1]Delft University of Technology

Variations in radar backscatter reflect changes in soil moisture, vegetation structure and the moisture content of the vegetation constituents. Radar's ability to "look into" the vegetation, can therefore provide unique insight into how water is pumped from the soil to the atmosphere through the vegetation, and how this pumping is regulated by the vegetation's physiological response to environmental conditions. Assimilation of radar data provides a way to estimate the states, and constrain the parameters related to carbon and water exchanges between the land surface and atmosphere.

The challenge in assimilating radar data into land surface models lies in the measurement operator. Advances in the representation of plant hydraulics in land surface models is improving our ability to simulate vegetation water dynamics. However, models used to simulate backscatter from vegetated surfaces still treat vegetation as a static dielectric medium. Furthermore, radiative transfer models require the dielectric properties of the soil and vegetation as well as a description of the size, shape, orientation and distribution of scatterers in the canopy. These parameters are seldom known, and they are not simulated in land surface models. Therefore, there is a mismatch between the representation of vegetation water dynamics in land surface models and our ability to simulate the backscatter variations due to those dynamics. Here, we will present preliminary results from a recent study where the use of machine learning is explored to address this mismatch.

June 16, 2021: Challenges in dealing with observations
9:35 AM EDT

Using the International Tree-Ring Data Bank (ITRDB) records as century-long benchmarks for land-surface models

Jina Jeong[1], Jonathan Barichivich[2,3], Philippe Peylin[2], Vanessa Haverd[†,4], Matthew J. McGrath[2], Nicolas Vuichard[2], Michael N. Evans[5], Flurin Babst[6,7,8], and Sebastiaan Luyssaert[1]

[1]Department of Ecological Sciences, VU University, 1081HV Amsterdam, the Netherlands.
[2]Laboratoire des Sciences du Climat et de l'Environnement, IPSL, CNRS/CEA/UVSQ, 91191 Gif sur Yvette, France.
[3]Instituto de Conservación Biodiversidad y Territorio, Universidad Austral de Chile, 5090000 Valdivia, Chile.
[4]CSIRO Oceans and Atmosphere, Canberra, 2601, Australia.
[5]Department of Geology ESSIC, University of Maryland, MD 20742-4211, USA.
[6]Dendro Sciences Group, Swiss Federal Research Institute WSL, Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland.
[7]School of Natural Resources and the Environment, University of Arizona, Tucson, USA.
[8]Laboratory of Tree-Ring Research, University of Arizona, Tucson, USA.
[†]Deceased 29 January, 2021.

The search for a long-term benchmark for land-surface models (LSM) has brought tree-ring data to the attention of the land-surface community as they have recorded growth well before human-induced environmental changes became important. The most comprehensive archive of publicly shared tree-ring data is the International Tree-ring Data Bank (ITRDB). Many records in the ITRDB have, however, been collected almost exclusively with a view on maximizing an environmental target signal (e.g. climate), which has resulted in a biased representation of forested sites and landscapes and thus limits its use as a data source for benchmarking. This study aims to propose advances in data processing to enable the land-surface community to re-use the ITRDB data as a much-needed century-long benchmark. Given that tree-ring width is largely explained by phenology, tree size, and climate sensitivity, LSMs that intend to use it as a benchmark should at least simulate tree phenology, size-dependent growth, differently-sized trees within a stand, and responses to changes in temperature, precipitation and atmospheric, CO2 concentrations. Yet, even if LSMs were capable of accurately simulating tree ring width, sampling biases in the ITRDB need to be accounted for. This study proposes two solutions: exploiting the observation that the variation due to size-related growth by far exceeds the variation due to environmental changes, and simulating a size-structured population of trees. Combining the proposed advances in modeling and data processing resulted in four complementary benchmarks - reflecting different usage of the information contained in the ITRDB - each described by two metrics rooted in statistics that quantify the performance of the benchmark. Although the proposed benchmarks are unlikely to be precise, they advance the field by providing a much-needed large-scale constraint on changes in the simulated maximum tree diameter and annual growth increment for the transition from pre-industrial to present-day environmental conditions over the past century. Hence, the proposed benchmarks open up new ways of exploring the ITRDB archive, stimulate the dendrochronological community to refine its sampling protocols to produce new and spatially unbiased tree-ring networks, and help the modeling community to move beyond the short-term benchmarking of LSM.

June 16, 2021: Challenges in dealing with observations
9:55 AM EDT

Advances and challenges for using paleoecological data for state data assimilation within a forest gap model

Ann Raiho[1]

[1]Colorado State University

Paleoecological data can be used to constrain forest gap model state variables and improve long-term predictions of forest demography. But, applying these constraints can be difficult because forest gap models can have high process variance and paleoecological data are often non-normally distributed. We built a new state data assimilation algorithm to estimate process variance and account for data containing zeros (left-censored) named the Tobit-Wishart Ensemble Filter (TWEnF). We applied our algorithm to a forest gap model at Harvard Forest using both tree-ring-derived species biomass and fossil pollen-derived fractional composition. We found that the TWEnF greatly improved predictions from the forest gap model and opened the door for further development. Possible next challenges include eliminating MCMC with Gibbs sampling methods, expanding paleoecological state data assimilation to larger spatial domains and models, updating parameters as well as states, and accounting for sudden events like disturbance.

June 16, 2021: Challenges in dealing with observations
10:10 AM EDT

Time scales in land data assimilation

Clara Draper[1]

[1]NOAA, Earth System Research Laboratories, Physical Sciences Laboratory

Land surface dynamics typically occur over longer time scales than those of the atmosphere. However, the observations available for assimilation are often of variables close to the land/atmosphere interface (e.g., near-surface soil moisture), which have faster time scales than the land surface variables of most interest (e.g., root-zone soil moisture). This presentation will review previously published work exploring aspects of these time scale differences. Draper and Reichle, (2015; HESS) assimilated AMSR-E near-surface soil moisture observations into that Catchment land surface model, showing that the assimilation improved the model near-surface soil moisture at both sub-seasonal (fast) and inter-annual (slow) time scales. More recent synthetic twin data assimilation experiments with a simple force-restore soil moisture model demonstrate that data assimilation algorithms can also effectively filter short time-scale observation errors to improve the longer-time scales of the root-zone soil moisture. This result was quite robust, and occurred even in the presence of large systematic model errors, and large fast-time scale observation errors. However, Draper and Reichle (2015) point out that the differences in time scales between the modeled and observed soil moisture have under-appreciated consequences in terms of soil moisture bias correction strategies, and in terms of the evaluation strategies used to assess different soil moisture estimates.

June 16, 2021: Challenges in dealing with observations
10:25 AM EDT

Technical challenges of assimilating observations with large spatiotemporal resolutions

[1]**Manuela Girotto**

[1]UC Berkeley

The GRACE and GRACE-FO missions (hereafter both will be referred to as “GRACE”) revolutionized large-scale remote sensing of the Earth’s hydrologic cycle, allowing scientists to peer deeper into the terrestrial water storage (TWS) across regional and continental scales. However, the spatial/temporal resolution and latency of GRACE data have limited their real time application in monitoring and characterizing land surface processes. In particular, because of the uniqueness of the spatial and temporal resolutions of GRACE data, it is not trivial to assimilate GRACE data and merge them with any other remotely sensed observations. This presentation will focus on benefits, challenges and applications of recent land surface data assimilation research efforts targeted at improving soil moisture, groundwater, and terrestrial water storage hydrological states using gravimetry observations from the GRACE and GRACE-FO missions.