120.081 Climate and Environmental Remote Sensing (VU, 2019S) – Exercise 3: Ecosystem model-data integration

Wouter Dorigo, Matthias Forkel, and Leander Mösinger

Climate and Environmental Remote Sensing Group, Department of Geodesy and Geoinformation, Technische Universität Wien, Gusshausstraße 27-29, 1040 Vienna, Austria

Correspondence to: Matthias Forkel (matthias.forkel@geo.tuwien.ac.at, for general questions regarding the exercise), Leander Mösinger (leander.moesinger@geo.tuwien.ac.at, for Python-related questions)

1 Introduction

Dynamic global vegetation models (DGVMs) are used to quantify past, present, and future dynamics of global ecosystems such as vegetation distribution, carbon and water cycling, and land management. Widely used is the LPJ (Lund-Potsdam-Jena) DGVM (Sitch et al., 2003) that in its most recent version (LPJmL4) represents carbon, water and energy fluxes in natural and managed vegetation and includes improved representations of permafrost and soil hydrology, fire, phenology, irrigation and water reservoirs, crop types and bioenergy plants (Schaphoff et al., 2018a) (Figure 1). LPJmL has been evaluated against various satellite and in-situ observations (Schaphoff et al., 2018b). Moreover, satellite observations have been used to estimate model parameters that govern the simulated phenology and gross primary production (Forkel et al., 2014).
The aim of this exercise is to analyse and evaluate LPJmL model results against satellite datasets of the fraction of absorbed photosynthetic active radiation (FAPAR), sun-induced fluorescence (SIF), and surface soil moisture (SSM) for some selected grid cells. Furthermore, the sensitivity of model parameters will be assessed for a grid cell in a Savannah ecosystem and an optimal model parameter set will be estimated.

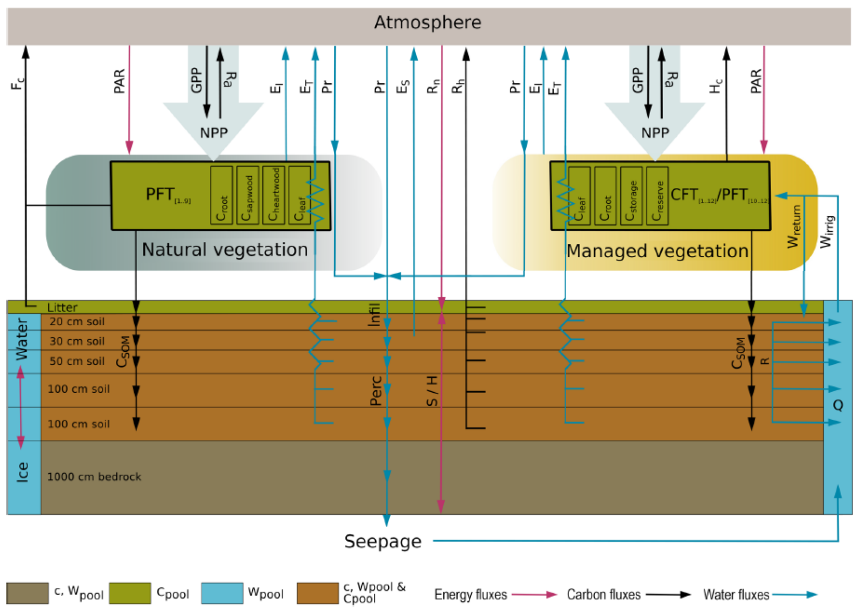

Figure 1. LPJmL4 scheme for carbon, water, and energy fluxes represented by the model. C – carbon; W – water; S – sensible heat conduction; H – latent heat convection; c – energy conduction; R_n – net downward radiation (input); PAR – photosynthetic active radiation; E_I – interception; E_T – transpiration; E_S – evaporation; Infil – infiltration; Perc – percolation; Pr – precipitation (input); GPP – gross primary production; NPP – net primary production; R_a – autotrophic respiration; R_h – heterotrophic respiration; H_c – carbon harvested; F_c – carbon emitted by fire; SOM – soil organic matter; R – run-off; Q – discharge.

Figure 1: Structure of the LPJmL4 model (Schaphoff et al., 2018a).

2 Data and methods

2.1 Datasets

Two datasets are provided for this exercise:

- LPJmL results from five grid cells (0.5° x 0.5° resolution) of a standard model run (cells in the Sahel, Mexico, Canada, Spain, and the Taymyr peninsula, Table 1). Model results are carbon and water fluxes, biomass, vegetation cover, soil moisture and FAPAR (Table 2). Additionally, model results for 50 different parameter sets are provided for the Sahel grid cell:
    - The folders "cell_32785/pars1" to "pars50" contain results from 50 model runs using different parameter sets.
    - The parameter sets and values are provided in the file "cell_32785_parameter-sets.txt".
- Satellite datasets:
    - FAPAR is provided from MODIS (MOD15A2H product) (Myneni et al., 2015, p.2)
    - SIF is provided from the GlobFluo algorithm from GOME-2 observations (Köhler et al., 2015)
    - SSM is provided from the ESA CCI surface soil moisture (SSM) dataset (Dorigo et al., 2017).

Table 1: Description of sites or grid cells for which LPJmL model results and observational data are provided.

<table>
<tr><th>Site name</th><th>Grid cell ID</th><th>Longitude</th><th>Latitude</th><th>Cropland area (%)</th><th>Model forcing</th></tr>
<tr><td>Sahel</td><td>32785</td><td>23.75°E</td><td>7.75°N</td><td>11</td><td rowspan="5">LPJmL has been driven with the Princeton Global Meteorological Forcing Dataset V2 at 0.5° resolution. The coverage of agricultural land and burned area was prescribed from observational datasets.</td></tr>
<tr><td>Mexico</td><td>6037</td><td>105.25°W</td><td>28.75°N</td><td>85</td></tr>
<tr><td>Canada</td><td>7460</td><td>99.25°W</td><td>52.25°N</td><td>0.3</td></tr>
<tr><td>Spain</td><td>25368</td><td>3.25°W</td><td>39.75°N</td><td>77</td></tr>
<tr><td>Taymyr</td><td>53569</td><td>102.25°E</td><td>75.75°N</td><td>0</td></tr>
</table>

2.2 Model parameters

LPJmL has a large number of model parameters (Schaphoff et al., 2018a). Two model parameters were selected for this exercise to investigate the sensitivity to FAPAR, SIF and soil moisture. The selected model parameters are (Schaphoff et al., 2018a):

- EMAX is the daily maximum evapotranspiration rate (mm day-1) (equation 112 in Schaphoff et al. 2018a). A value of 7 mm was used for the tropical broadleaved rain-green plant functional type in the standard model run.
- WATER_BASE is the water availability at which half of the canopy is developed (denoted as b_x in equations 41 and 42 in Schaphoff et al. 2018a). The value of 22% was used for the tropical broadleaved rain-green plant functional type in the standard model run. A higher value should lead to later leaf development in water-limited ecosystems.

2.3 Groups, data access, and software

- **TUWEL is used to organize the exercises (https://tuwel.tuwien.ac.at/course/view.php?id=18244).**
- The exercise should be done in the same groups as in exercises 1 and 2.
- Download the data and example scripts from **TUWEL (climers_exercise03.zip**). The package also contains some supplementary material.

We recommend doing the exercise in Python. You are free to use other software but we provide support only for Python. Information on how to install and use miniconda for this exercise on Windows and Linux systems is provided in the Appendix (below). Please use the environment file **climers_env_ex01.yml from exercise 1** to get started with your own analysis. The file **exercise03.py** contains a reader function to read all csv-files for this exercise. Please contact Leander Mösinger in case of Python-related questions.

Table 2: Description of LPJmL model outputs.

Variable / file name	Description	Unit
fpc	Annual Foliar Projective Cover, i.e. the fractional coverage of plant functional types. Column names or plant functional types in the files are: • NatStand = fraction of natural vegetation in a grid cell. The coverage of agricultural area is 1 – NatStand and was prescribed to LPJmL4 by an input dataset. The coverage of natural vegetation types is simulated by LPJmL4: • TrBE = tropical broad-leaved evergreen tree • TrBR = tropical broad-leaved rain-green tree • TeNE = temperate needle-leaved evergreen tree • TeBE = temperate broad-leaved evergreen tree • TeBS = temperate broad-leaved summer-green tree • BoNE = boreal needle-leaved evergreen tree • BoBS = boreal broad-leaved summer-green tree • BoNS = boreal needle-leaved summer-green tree • TrH = tropical herbaceous vegetation (grasses) • TeH = temperate herbaceous vegetation (grasses) • PoH = polar herbaceous vegetation (tundra)	Fraction per grid cell
mburnt_area	Monthly burned area was described to the fire module of LPJmL using a satellite-observed datasets (after 1996) (Giglio et al., 2013). Before 1996, the observed burned area was resampled according to a ranking of hot and dry years.	ha month-1
mevap	Monthly total evaporation	mm month-1
mfapar	Monthly average FAPAR	fraction
mfirec	Monthly total fire carbon emissions	gC m-2 month-1
mgpp	Monthly total gross primary production	gC m-2 month-1
minterc	Monthly total interception	mm month-1
mnpp	Monthly total net primary production	gC m-2 month-1
mrh	Monthly total heterotrophic respiration	gC m-2 month-1
mswc1	Monthly average soil water content (fraction of saturation) in upper soil layer (0-20 cm depth)	fraction
mtransp	Monthly total transpiration	mm month-1
vegc	Annual vegetation carbon content (above- and belowground biomass)	gC m-2

2.4 Tasks

2.4.1 Task A: Vegetation dynamics in different ecosystems

Use the LPJmL results from the standard model run for the five grid cells.

- Compute monthly evapotranspiration (ET) and net ecosystem exchange (NEE).
- Plot for each grid cell monthly time series of FAPAR, soil water content (SWC), ET, GPP, and NEE.
- Compute annual total gross primary production (GPP) and annual net biome productivity (NBP).
- Plot annual time series of GPP, NBP, biomass and of the foliar projective cover (FPC) of the dominant tree and grass plant functional types.

Describe and discuss the simulated vegetation dynamics for each ecosystem.

2.4.2 Task B: Evaluation of simulated FAPAR, GPP and SWC

Evaluate for each grid cell the simulated FAPAR, GPP and SWC against the satellite datasets.

- Select five model performances measures from Janssen and Heuberger (1995) to quantify the model performance for FAPAR, SIF and SSM. Try to select model performance measures that quantify different aspects of model-data agreement (e.g. bias, variance, correlation, error, overall performance).

- *Hint 1:* SIF is linearly related to GPP (Frankenberg and Berry, 2018). Despite the linear relation between SIF and GPP, please note that those are different ecophysiological variables with different units. Hence, you cannot compare SIF and GPP in absolute terms/units but you can use SIF to evaluate the temporal dynamic of GPP.
- *Hint 2:* Please note that the range of the SWC from LPJmL is not comparable with the range of ESA CCI SSM, which depends on the retrieval algorithm for satellite SSM. Hence, you need to normalize one or the other before the computation of model performance measures.

Visualize the evaluation results and discuss (a) differences between model performance metrics and (b) the general model performance. Compare the model performance for FAPAR and GPP/SIF and describe the relations between both variables.

2.4.3 Task C: Model parameter sensitivity – towards model optimization

Use the LPJmL results for the different parameter sets for the Sahel grid cell (cell_32785).

- Select from task B two model performance metrics and compute the performance for FAPAR, SIF and SWC for each parameter set. Visualize the dependency between model parameters and model performance. Can you identify optimal parameter values?
- Compute from the results for each parameter set mean annual ET, GPP, NBP, biomass, and FPC of the tropical broad-leaved rain-green tree. How do they change dependent on the model parameterisation?
- Imagine that you plan to optimize model parameters of LPJmL jointly against satellite data of FAPAR, SIF and SSM. Can you identify a subset from the parameter sets that give optimal results for these variables? How would you make sure within such an optimization that the other variables stay within plausible ranges?

3 Expected results and examination procedure

- Summarize the results of your group in a **report**. The report should be written in the style of a scientific publication (research article). Table 3 lists the requirements for the report. It is especially important that you *describe the contribution of each group member* to the report (Table 3).
- Present a preliminary version of the report in an **examination meeting** with the lecturers **(24.-28. June)**. Scheduling of the meetings will be done though TUWEL. Optionally, you can present the exercise in a PowerPoint presentation. During the examination meeting, every group member will get **additional questions** related to the contents of the lecture and the corresponding chapters and review questions in the extra literature provided.
- **Upload the final version of your report** (.pdf) at TUWEL as a group at the latest until **4. July 2019.**
- Every group member will receive a **mark** for (a) the specific contribution to the report (60%), and (b) for the quality of responses to questions during the examination meeting (40%).

Table 3: Requirements for the report of exercise 2.

Item	Description
Template	Please use this description of the exercise as template for your report. The template is largely based on the template of the Copernicus journals from the European Geosciences Union. If you prefer to use Latex, please download the Copernicus Publications LaTeX package from https://www.biogeosciences.net/for_authors/manuscript_preparation.html
Chapters	Include the following chapters in your report: 1. Introduction 2. Data and methods 3. Results and Discussions (can be also split in 3. Results and 4. Discussions) 4. Conclusions References

	Appendix (optional). The appendix can contain data tables, short descriptions of own developed code or functions (if no standard packages were used), or other material. Please do not include additional figures in the appendix.
Length	The report should have **at least 1000 but no more than 3000 words** (chapters 1 – 4, i.e. word limits are without title, author names/email, references, and appendix).
Figures and tables	Present the results as **figures** and **tables** in your report. Please follow as much as possible these recommendations for the design of your figures: https://www.biogeosciences.net/for_authors/manuscript_preparation.html
References	We do not expect that you do an extensive literature research for this exercise. However, please include citations and references for the used data and methods. Additionally, please include references if you refer to any publications in your report. Use the citation style of EGU Copernicus publications: https://www.biogeosciences.net/for_authors/manuscript_preparation.html If you are using Zotero for reference management, you can get the reference style template from: https://www.zotero.org/styles/biogeosciences
Author contributions	Add at the end of your report a section that describes the contribution of all group members. Example: *X* analysed the data for task A and B. *Y* analysed the data for task C. *Z* mainly wrote the report with inputs from *X* and *Y*. It is your own responsibility to fairly distribute the work within your group.

References

Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Chung, D., Ertl, M., Forkel, M., Gruber, A., Haas, E., Hamer, P. D., Hirschi, M., Ikonen, J., de Jeu, R., Kidd, R., Lahoz, W., Liu, Y. Y., Miralles, D., Mistelbauer, T., Nicolai-Shaw, N., Parinussa, R., Pratola, C., Reimer, C., van der Schalie, R., Seneviratne, S. I., Smolander, T. and Lecomte, P.: ESA CCI Soil Moisture for improved Earth system understanding: State-of-the art and future directions, Remote Sens. Environ., 203, 185–215, doi:10.1016/j.rse.2017.07.001, 2017.

Forkel, M., Carvalhais, N., Schaphoff, S., v. Bloh, W., Migliavacca, M., Thurner, M. and Thonicke, K.: Identifying environmental controls on vegetation greenness phenology through model–data integration, Biogeosciences, 11(23), 7025–7050, doi:10.5194/bg-11-7025-2014, 2014.

Frankenberg, C. and Berry, J.: Solar Induced Chlorophyll Fluorescence: Origins, Relation to Photosynthesis and Retrieval, in Comprehensive Remote Sensing, pp. 143–162, Elsevier., 2018.

Giglio, L., Randerson, J. T. and van der Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), J. Geophys. Res. Biogeosciences, 118(1), 317–328, doi:10.1002/jgrg.20042, 2013.

Janssen, P. H. M. and Heuberger, P. S. C.: Calibration of process-oriented models, Ecol. Model., 83(1–2), 55–66, doi:10.1016/0304-3800(95)00084-9, 1995.

Köhler, P., Guanter, L. and Joiner, J.: A linear method for the retrieval of sun-induced chlorophyll fluorescence from GOME-2 and SCIAMACHY data, Atmospheric Meas. Tech., 8(6), 2589–2608, doi:https://doi.org/10.5194/amt-8-2589-2015, 2015.

Myneni, R. B., Knyazikhin, Y. and Park, T.: MOD15A2 MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 1km SIN Grid, Boston University and MODAPS SIPS, NASA, [online] Available from: https://doi.org/10.5067/MODIS/MOD15A2.006 (Accessed 11 January 2019), 2015.

Schaphoff, S., von Bloh, W., Rammig, A., Thonicke, K., Biemans, H., Forkel, M., Gerten, D., Heinke, J., Jägermeyr, J., Knauer, J., Langerwisch, F., Lucht, W., Müller, C., Rolinski, S. and Waha, K.: LPJmL4 – a dynamic global vegetation model with managed land – Part 1: Model description, Geosci Model Dev, 11(4), 1343–1375, doi:10.5194/gmd-11-1343-2018, 2018a.

Schaphoff, S., Forkel, M., Müller, C., Knauer, J., von Bloh, W., Gerten, D., Jägermeyr, J., Lucht, W., Rammig, A., Thonicke, K. and Waha, K.: LPJmL4 – a dynamic global vegetation model with managed land – Part 2: Model evaluation, Geosci Model Dev, 11(4), 1377–1403, doi:10.5194/gmd-11-1377-2018, 2018b.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K. and Venevsky, S.: Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, Glob. Change Biol., 9, 161–185, 2003.