Tethys-Chloris (T&C): model setup, validation, and sensitivity analysis

Material:

- **Source Code and Case Study Materials**:
  - T&C source code
  - Scripts for meteorological data pre-processing
  - *prova* and *parameter* files for the case studies

- **Input Data:**
  - Files containing CO_2 air concentration

- **Reference Materials:**
  - Parameter Lookup File: *T&C_Variables_LIST_PlotScale.pdf*
  - Technical Reference: *T&C model technical reference.pdf*

- **Visualization and Validation Tools:**
  - *GRAPHY_MOD* file for results visualization
  - *validation_X* file for model validation.

- **Supporting Data:**
  - Flux tower data from case studies (instructions for their download provided below).

Objectives. This session is designed to help students learn how to prepare input files and run a basic case study using the plot-scale version of the T&C model (*Fatichi et al., 2012*). The document outlines the key steps, including pre-processing input data, verifying critical model settings, understanding model inputs and outputs, and conducting model validation and sensitivity analysis. The tutorial here will focus on two case studies from Switzerland (Chamau and Laegeren).

Part A: input preparation, model setup and validation

1. MATLAB add-ons

Please make sure you have downloaded MATLAB on your PC. To run T&C smoothly, the following Matlab toolboxes must be installed:

- Financial Toolbox
- Global Optimization Toolbox
- Image Processing Toolbox
- Optimization Toolbox
- Parallel Computing Toolbox
- Partial Differential Equation Toolbox
- Statistics and Machine Learning Toolbox
- Symbolic Math Toolbox

2. Prepare the meteorological input file

Meteorological input data, such as air temperature, precipitation, and solar radiation, are essential inputs to the ecohydrology model T&C. To prepare the input data required for T&C, we will use data from the FLUXNET monitoring network (https://fluxnet.org/). Particularly, the Swiss FluxNet**,** a regional subset of FLUXNET, currently encompasses six long-term ecosystem monitoring sites in Switzerland.

- To prepare the meteorological input data, start by visiting the Swiss FluxNet website.
- Select our first case study site, **Chamau**, and download the corresponding data.
- After unzipping the folder, the file **'FLX_CH-Cha_FLUXNET2015_FULLSET_HH_2005-2023_1-3.csv'** will be used as input for the model (and for validation later on).
- Open the file **'read_fluxnet2015_tandc_Teaching.m'** under the folder Pre_process and set up your working path configuration along with the geographic information of Chamau. Information for all FLUXNET sites can be accessed at fluxnet.org (you can find the description of FLUXNET data at this page in the website: FluxNet2015 Data).

This is a brief summary of the main steps for FLUXNET2015 data cleaning and processing implemented in the *read_fluxnet2015_tandc_Teaching.m* file:

- User-Defined Variables and Configuration: defines required variables to extract from FLUXNET data, including temperature (TA_F), precipitation (P_F), radiation (SW_IN_F, LW_IN_F), wind speed (WS_F), etc.
- Read in the FLUXNET site list with all the metadata: manually specify site metadata, including longitude, latitude, elevation, and time zone, for later time conversions.
- Find the index of the required variables in the raw data.
- Aggregating half-hourly data to hourly data: sums precipitation data while computing the mean for other variables.
- Variable screening and physical constraints: applies physical constraints to filter out unrealistic values, handling missing values by replacing them with the mean values of the variable at the same Julian day and hour across different years.

- Computing additional variables: calculates, e.g., vapor pressure at saturation (e_{sat}) and vapor pressure (e_a).
- Saving cleaned data: saves the final set of input variables required for running the T&C model, as well as the file for model validation.

Then:

- Run the script: the files **'Data_X_run.mat'** and '**Res_X.mat'** will be generated. To understand the meaning of each variable in these files, refer to the **T&C_Variables_LIST_PlotScale.pdf** or the lecture slides. **'Data_X_run.mat'** contains the necessary meteorological forcing to run T&C, while '**Res_X.mat'** contains some variables for model validation (see point 5 below).
- Repeat this process for the second study site, **Laegeren**.

Once the files are ready, consider the following questions:

- *What meteorological data is extracted from the file as input and for validation?*
- *How different are the two case studies in terms of climatic conditions? To what biomes do they belong according to the Whittaker classification (Whittaker, 1970)?*

3. Understand the case studies

Let's begin with a simple case study in Chamau**.**

- Open the file **'prova'** and ensure that your MATLAB working directory is set to the same path as the file (Figure 1 below).

The command `load(Data_CH-Cha_run.mat)` loads the meteorological data for Chamau (make sure to put the metrological file in the same folder as 'prova'). Before running the simulation (point 4 below), take a moment to familiarize yourself with the case study and answer the following questions:

- *What are the mean annual precipitation and mean temperature in Chamau?*
- *What is the vegetation type and landcover in Chamau? Is there any land management?*
- *What temporal patterns do you expect in Leaf Area Index (LAI) and Gross Primary Productivity (GPP)?*
- *What are the soil depth and soil texture? How many soil layers are set?*

Hint: Chamau is an intensively managed grassland. In the **'MOD_PARAM'** file, check the variables **ZR95_H**, **ZR95_L**, and **Ccrown**. Chamau is covered exclusively with low vegetation (grass), and the cutting dates for the grass are indicated by **Mpar_L(1).jDay_cut**. The abbreviations and meanings of these variables can be found in **T&C_Variables_LIST_PlotScale.pdf**.

Remember to perform the same checks before running the simulation for a new case study. For example, try answering the same questions for Laegeren. While for the case study in Chamau all parameters were already provided, some key plant parameters and soil properties are not directly provided for the Laegeren case study. Please follow the steps below to search for specific site information and setup T&C in Laegeren.

To set up the model in Laegeren:

- Open the **'MOD_PARAM_LAEGEREN'** file and provide the require soil texture information for Laegeren. You can consider the soil type to be a sandy loam and refer to the soil texture triangle to get reasonable textural values.
- Vegetation cover in Laegeren is simplified to two types, beech and fir (as defined in Ccrown=[0.55 0.45]). In the vegetation parameters definition in T&C, if more than one vegetation cover is set, the parameter will have more than one element (one for each vegetation type – the indexing follows that of the vegetation type). For example, KnitH=[0.35 0.20] means the canopy nitrogen decay coefficient of the first vegetation type is 0.35 and of the second vegetation type is 0.20. Usually, vegetation parameters are taken from the existing literature and tuned in the model calibration phase. In the file, some parameters are currently missing (set as *NaN* for now) for beech in Laegeren: provide them based on the listed parameter values in the table below (Table 1). Note that only parameters listed in Table 1 need to be filled.

Table 1. Look-up table for some vegetation parameters

Parameter /Location	Chamau	BAY	Davos	Hainich	HARVARD	Laegeren
Species	C3 Grassland Managed and smooth meadow grass,	H: Evergreen Forest [Norway Spruce].	H: Evergreen Forest [Norway spruce]	H: Deciduous forest [Beech, ash, and maple]	H: Deciduous Forest [red oak and red maple]	H: Deciduous Forest [Beech] (Para only for deciduous)
aSE	2	0	0	1	1	
Sp_LAI_In	0.2	0.1	0.1	0.15	0.2	
d_leaf	0.8	0.25	0.25	4	4	
Ha	55	72	72	76	76	
Tcold	-2	-20	-20	3	5.5	
Tlo	0	8.5	4.5	5	5.5	
Vmax	96	40	44	76	62	

After the missing parameters are set in Laegeren, the model parameter file is ready. Similar steps can be performed to implement T&C in a new case study. Note however that, while parameters here were provided in a Table for simplicity, in general these should be extracted from existing databases – e.g. SoilGrids, TRY Plant Trait Database – or from the literature. This may be the case in your group project.

4. Ready to run the simulation

- Ensure that you are in the correct working directory. Check if the **'prova'** file is visible in the current folder window of MATLAB. If not, right-click on the file name in the editor and select **"Change Folder"** to set the correct working directory. Additionally, confirm that the **'Inputs'** and **'T_C_Code'** folders are located in the correct directory, as shown in Figure 1 below.

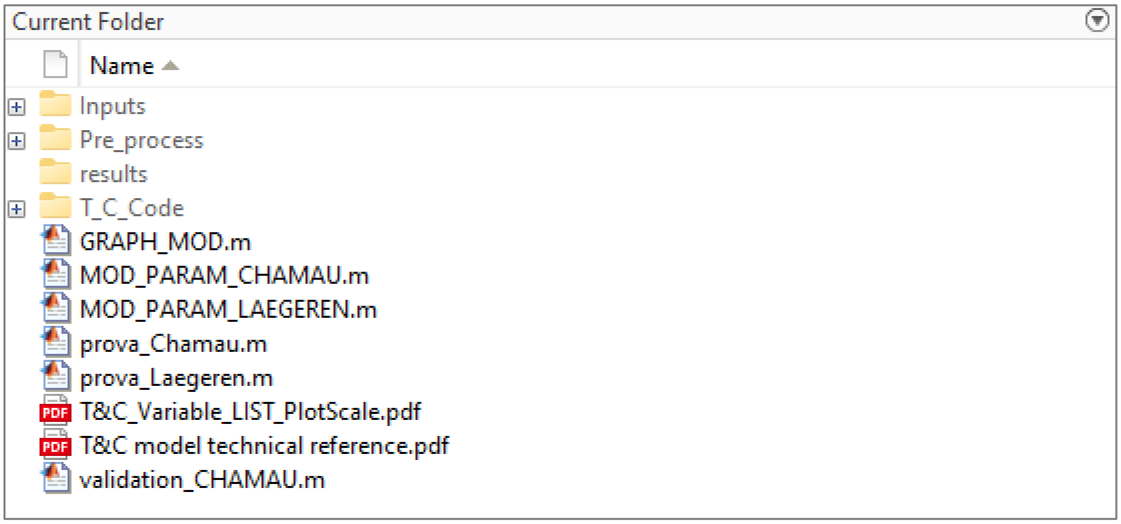

Figure 1. Correct working path

- Next, open the **'prova'** file and set up the simulation period. Verify the measurement period of your meteorological data in the **'Date'** variable. You can use the **datevec()** function to convert date numbers into the format yyyy-mm-dd-hh-(ss) to identify the available meteorological data period. Note that the CO_2 data is loaded from the 'Ca_Data.mat', which is only available until the end of 2022.

  - *What is the time period covered by the meteorological data?*

```
%%%%%%%%%%%%%%%%%%%%% make sure you understand your data availability
dateNum1 = datenum(2005, 1, 1, 0, 0, 0);
dateNum2 = datenum(2014, 12, 31, 23, 0, 0);
index1 = find(Date >= dateNum1, 1, 'first');
index2 = find(Date >= dateNum2, 1, 'first');
```

Figure 2. An example of simulation period setup.

- Set **dateNum1** and **dateNum2** according to the provided meteorological data period (Figure 2). Then, review the settings of **PARAM_IC** and **Directory** in the script. Investigate how these variables are defined and understand their purpose. Additionally, explore the function **filesep** in MATLAB, which is used to insert the appropriate file separator (/ or \) based on the operating system, ensuring compatibility across platforms.

- Once everything is set up, run the simulation for the case study in Chamau (and/or Laegeren) using the provided meteorological input data and default parameters. After starting the simulation, note the total runtime for the model.

```
COMPUTATIONAL TIME [h]
    0.0377

 COMPUTATIONAL TIME [ms/cycle]
    1.5504

   4.0018e-10

  -3.1592e-14
```

Figure 3. An example of running 10 years simulation in Chamau on a PC.

After the simulation is completed, you will see a message similar to that of Figure 3, all the variables are saved both in your MATLAB workspace and in the **'results'** folder. You can access and analyze these variables directly in MATLAB.

5. Result checking and model validation

To quickly check some results, you can easily plot variables. For example, try plotting **LAI_H** and **LAI_L** using the **plot** function. Observe what each plot reveals about the behavior of high and low vegetation LAI.

To get an overview of the results, use the provided **'GRAPH_MOD.m'** script. Simply type `GRAPH_MOD` in the Command Window. This will open a window and display several figures summarizing the model output. Detailed information about these figures can be found in the source code of **'GRAPH_MOD.m'**.

In the model validation phase, the goal is to compare the observed and simulated values of the extracted variables and assess how well the current model parameters perform. Data from the FluxNet towers in Chamau and/or Laegeren is used for model validation.

- Open the **'validation_X_incomplete'** file, which contains incomplete code to load data for validation. Refer to the introduction page of the FluxNet data description and answer the following:

- *What data is used for model validation in the script?*
- *What is the temporal resolution of these variables, both from measurements and from the model simulation?*

- *Plot the observed data (similar to Figures 4 and 5 below) and identify potential errors (such as missing values, clear outliers in the measurements, if any). If errors are found in the observed data, fix or discard the affected periods before proceeding with validation.*

- Start by plotting figures similar to Figures 4 and 5 below to compare observed and simulated data.

- To quantify the goodness of fit, use at least three statistical metrics, such as:

    - **Root Mean Square Error** (RMSE): measures the average magnitude of errors between observed and simulated values.
    - **Nash-Sutcliffe Efficiency** (NSE): evaluates the model's predictive skill compared to the mean of observed values.
    - **Coefficient of Determination** (R^2): Indicates the proportion of variance in the observed data explained by the model.

When interpreting the results, note that an R^2 value of 0.4 is already considered satisfactory for daily GPP validation (for the application here, it is not necessary to achieve results as good as those in the example figures below).

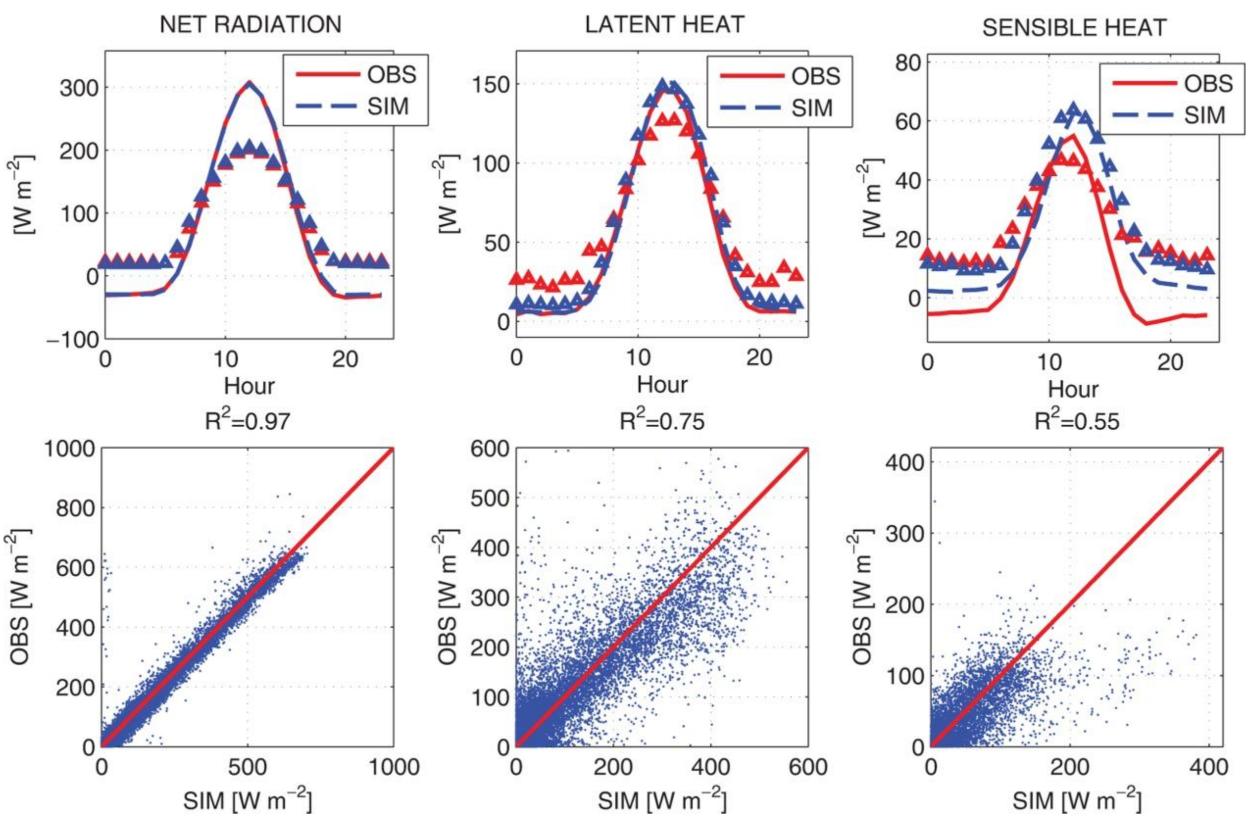

Figure 4. An example of results from Fatichi et al. (2014): comparison between the observed (OBS) and simulated (SIM) average daily cycles of net radiation, latent heat, and sensible heat for the location of Chamau. The triangles represent the standard deviations. Scatter plots with the determination coefficients R^2 are also shown.

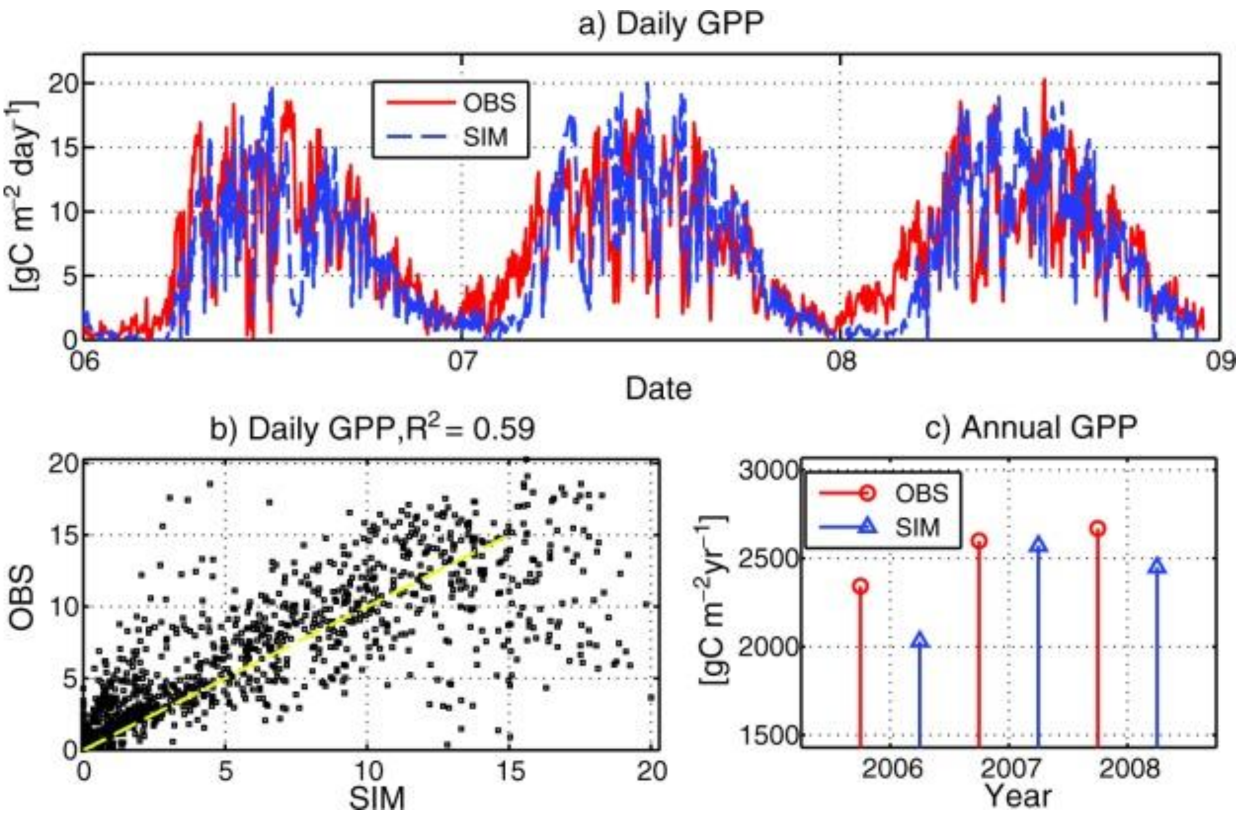

Figure 5. An example of results from Fatichi et al. (2014): comparison between the observed (OBS) and simulated (SIM) (a) daily and (c) annual GPP for the location of Chamau. A scatter plot with the determination coefficient R^2 is shown for (b) daily GPP.

- Focus on identifying discrepancies and calibrating the parameters to improve the fit iteratively. For the model calibration, focus primarily on calibrating the target variable GPP and try to identify the parameter set that achieves the highest goodness of fit. In the model, parameters ending with **_H** correspond to high vegetation, while those ending with **_L** pertain to low vegetation. Experiment with changing these values and evaluate their effects on GPP.
- Table 2 shows some potentially important parameters that influence the model output. Manually adjust these parameters and re-run the simulation to assess their impact. Try to get the combination of parameters that allows you to best match the data (model calibration). Note that several techniques exist for automatic model calibration, but this will not be explored here for the sake of time. Also, for complex models such as T&C the setup of an automatic calibration procedure may be quite complicated (given the very large parameter space).
- After adjusting the parameters, observe the changes in GPP and document their effects in Table 2. For example ,did the parameter changes increase or decrease GPP? Did they significantly affect the temporal patterns of GPP? Summarize the relationship between parameter changes and the simulation results in the table without including specific statistical metrics.

Table 2. Effect of variations in some key parameters on daily GPP/LAI dynamics.

Parameters	Unit	Meaning	Impact on GPP/LAI pattern
Tcold_H / Tcold_L	°C	Air temperature threshold for shedding of leaves, High/Low Vegetation	
Vmax_H / Vmax_L	μmol CO_2/ m^2 s	Maximum Rubisco capacity at 25°C leaf level, High/Low Vegetation	
Tlo_H / Tlo_L	°C	Threshold temperature for leaf onset, High/Low Vegetation	

Part B: Sensitivity analysis and scenarios

In this second part, we will perform a sensitivity analysis and build simple climate change scenarios.

1. Sensitivity analysis

Modify the soil/vegetation parameters listed below and explore how these parameters impact key ecohydrological variables (increase, decrease, in percentage). Here is an example in Laegeren.

Variables	GPP	ET (Latent hear, QE)	Leakage
Soil properties			
(Sand) Psan=88% Pcla=5%			
(Silt loam) Pan=10% Pcla=5%			
(Clay) Psan=25% Pcla=50%			
Vegetation properties			
Root depth decrease to 500 [mm]			
Root depth increase to 1200 [mm]			
Specific leaf area -50%			
Specific leaf area +50%			
Vmax = 50 [μmol CO_2/ m^2 s]			
Leaf carbon nitrogen ratio -50%			
Leaf carbon nitrogen ratio +50%			

2. Analysis of simple scenarios

Run some additional simulations changing the meteorological forcing and explore how the modified forcing impacts target ecohydrological variables (increase, decrease, in percentage).

Variables	GPP	ET (Latent heat, QE)	Leakage
Meteorological forcing			
Ta +3ºC			
Pr -50%			
CO_2 +300 ppm			

Find out more

Laboratory of Catchment Hydrology and Geomorphology: https://www.epfl.ch/labs/change/

T&C GitHub source code : https://github.com/simonefatichi/TeC_Source_Code

References

Fatichi, S., Ivanov, V. Y., & Caporali, E. (2012). A mechanistic ecohydrological model to investigate complex interactions in cold and warm water-controlled environments: 1. Theoretical framework and plot-scale analysis. *Journal of Advances in Modeling Earth Systems*, *4*(2).

Fatichi, S., Zeeman, M. J., Fuhrer, J., & Burlando, P. (2014). Ecohydrological effects of management on subalpine grasslands: From local to catchment scale. *Water Resources Research*, *50*(1), 148-164.

Whittaker, R. H. (1970). *Communities and ecosystems*.