Manual – PCLake

Manual with exercises and examples

NIOO

WAGENINGEN UR
For quality of life

stowa FOUNDATION FOR APPLIED WATER RESEARCH

PBL Netherlands Environmental Assessment Agency

Contents

PART 1: PREPARE INPUT

To run PCLake it is important to first look at your input data. You need to take several steps first to be able to run PCLake for your specific case. Moreover, preparing your data will also help you to systematically look at your system. This in itself can already provide you with valuable knowledge.

For this step, you can use the excel document called "InputPreparation_PersonalComputerLake.xlsx"

STEP 1: Prepare the water balance

The first step you take is to define the water balance. The water balance consists of all water flowing in or out of the lake. PClake has for different water inflow parameters:

Qin = [river inflow] + [precipitation] mm/d

Qout = [river outflow] mm/d

Qev = [evaporation] mm/d

cQinf = [Groundwater inflow] mm/d

Now, calculate the values in the white boxes by transferring them to the units of PCLake (mm/d).

Question: Is the estimated depth fluctuating? If yes how much?

Question: What is the estimated depth at the end of the year?

Question: What could happen with simulations if the water depth fluctuates too much?

STEP 2: Prepare the nutrient budget

The next step is to make a good nutrient budget. A nutrient budget for PCLake means to estimate all external input of phosphorus (TP) and nitrogen (TN) that comes into the lake (g/m2 lake area/d). Examples of external nutrient sources are:

- River inflow of nutrients
- Wastewater
- Run-off from surroundings
- Feeding birds
- Bird droppings
- Fishing bait
- Dog faces

Our example has both river inflow and baits from recreational fishing that forms an external nutrient load. Now, calculate the values in the white boxes by transferring them to the units of PCLake (g/m2/d).

Question: What is the average TN-Load? What is the average TP-Load?

Question: Do you think the lake will be turbid? Why?

Question: What is the N:P ratio? Does this mean this lake is nitrogen-limited, phosphorus-limited or co-limited?

Question: What might be other important nutrient sources not included in this calculation?

STEP 3: Other parameters

Temperature

The water temperature determines the growth speed of all kinds of organisms. PCLake has two kinds of water temperatures, that of the epilimnion (upper layer) and that of the hypolimnion (lower layer). In our example, we have a shallow lake that we assume has never been stratified. Therefore we only will look at the epilimnion temperature. The temperature of our example is already provided in the InputPreparation_PersonalComputerLake.xlsx file.

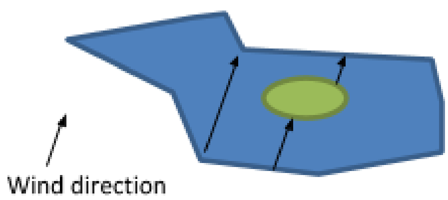
In case for your own lake, you might not have measured data available. Therefore you can use other models to estimate the temperature in your lake. For example, you can go to http://www.flake.igb-berlin.de.

Optional here: calculate the temperature of a lake with a depth of 2m, located at 41°53'07.3"N 76°21'06.2"W (our example), given the type of water quality estimated in STEP2 and a fetch calculated in STEP 1.

Optional question: Does this lake regularly stratify?

Fetch

The fetch is measured as the longest distance over water measured in the wind direction. This means that if the wind hits an island in the lake, the fetch is measured as the distance from the shore to the island. In the example below the longest arrow is the fetch of this lake. An approximation is the square root of the lake area.



Calculate the fetch of our example.

Question: is our example a large or small lake?

Sediment type

The sediment type is important e.g. to the binding of phosphorus. Clay and peat are known to strongly bind phosphorus for example. Moreover, sediment type determines the resuspension (e.g. sand is heavier so less likely to resuspend compared to clay). There are 5 parameters important for the sediment type:

- Clay content — flutum — gDWclay/gDW
- Iron content — fFeDIM — gFe/gDW
- Aluminium content — fAlDIM — gAl/gDW
- Organic matter content — fDOrgSoil — gOrg/gDW
- Dry Weight per gram sediment — fDTotS0 — gDW/gSed

Sometimes not all these parameters are available for your own lake. In that case, you can use the table provided in Janse 2005 (Janse, J.H., 2005. Model studies on the eutrophication of shallow lakes and ditches. Wageningen University) to estimate your parameters.

sediment type settings from Janse 2005					
type	dry matter fDTotS0	organic matter fDOrgSoil	lutum flutum	Fe fFeDIM	Al fAlDIM
clay	0.3	0.08	0.4	0.04	0.04
clay/peat	0.2	0.165	0.4	0.04	0.04
clay/sand	0.4	0.08	0.215	0.0215	0.0215
peat	0.1	0.25	0.4	0.04	0.04
sand	0.5	0.08	0.03	0.003	0.003

In our example, we assume a clay/peat lake. Look up the parameters for this type of sediment

Question: look at the different parameters. Do you understand the difference and their impact on water quality?

STEP 4: Collect all input values

Now it is time to summarize all the data you have. Put all values calculated in previous steps together and finalize the input for PCLake.

Question: Think for each value if you can consider this value as constant over the run.

STEP 5: Think of all switches

PCLake has numerous switches. These allow flexibility for the modeller to use the model. On the other hand, it makes PCLake also a complex model. Have a look at all the possible switches.

Tip: if you want to know the exact effect of a certain switch you can always dive into the equations and try to understand the effect.

Question: Switches are important in PCLake. Think for each switch whether you want it at 0 or 1

Question: Are the switches for Personal Computer Lake logical?

PART 2A: INSTALL PCLAKE, PCLake+ and PCDitch

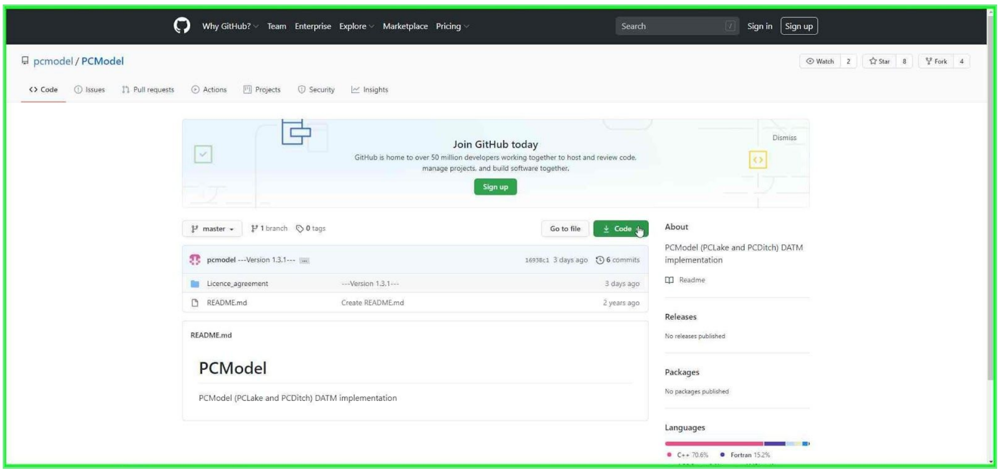
There are currently three models in the PCmodel family: PCLake, PCLake+ and PCDitch. PCLake is the old version of PCLake+ and was developed by Jan Janse. While PCLake is only available for shallow lakes, PCLake+ can simulate both shallow and deep lakes. Moreover, some errors in PCLake have been resolved in PClake+. PCDitch is a model for linear waters such as ditches and streams.

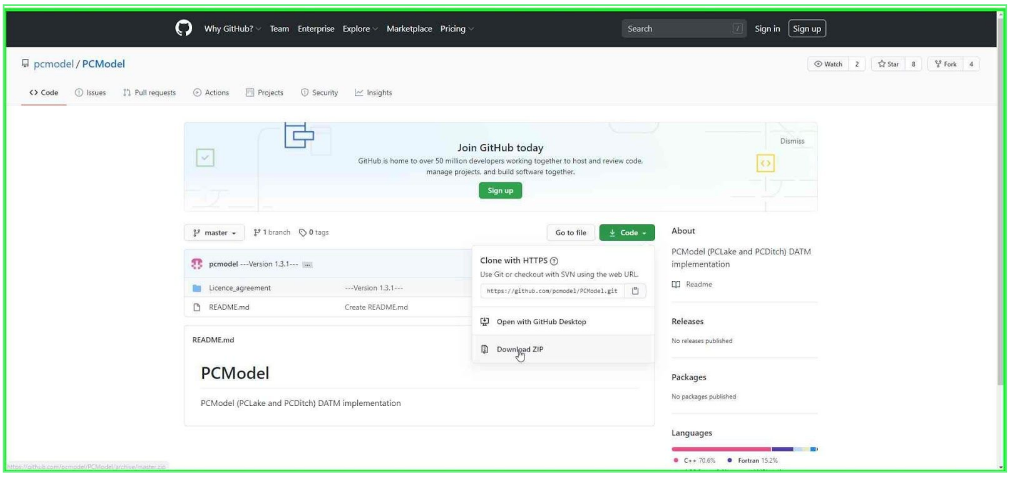
In this manual, we will focus on PCLake/PCLake+. PCDitch works similarly and is installed within the same package.

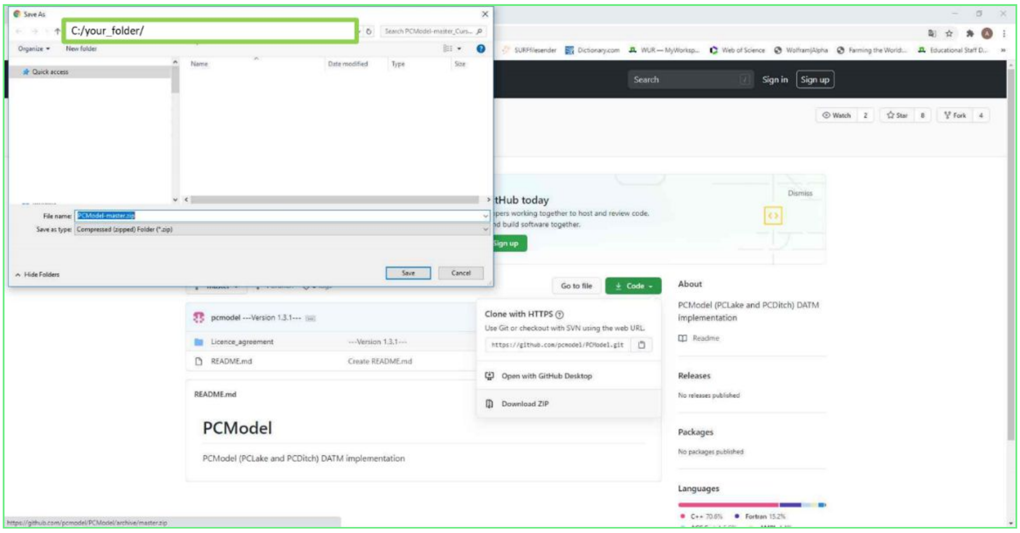
STEP 6: Download the model(s) and instal

To install PCLake, PCLake+ and PCDitch,

1) go to GitHub https://github.com/pcmodel/PCModel.
2) Click on ‘code’ --> download ZIP
3) Store the zip file at any location on your computer
4) Unzip the zip file at any location on your computer (preferred is a local hard-drive, other storage locations will make the software slow or even not working).
5) The package is stand-alone, and the directory tree should remain intact. So you can store it at any location, but do not change the folder structure of the file you just downloaded.





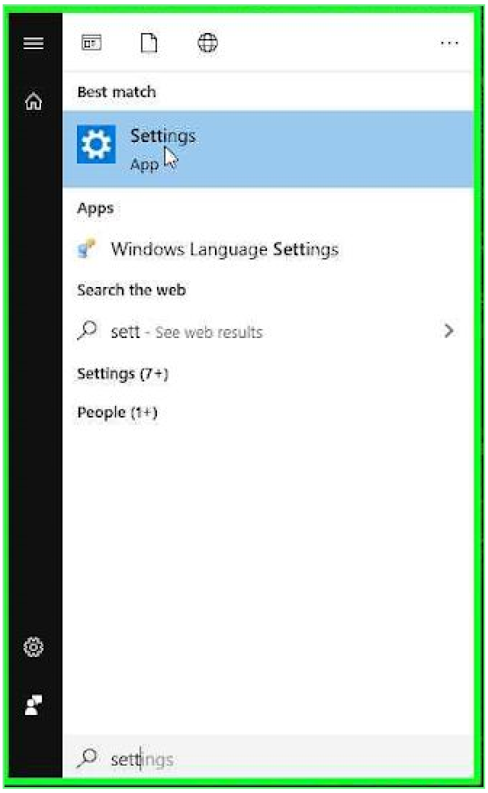


STEP 7: Computer settings

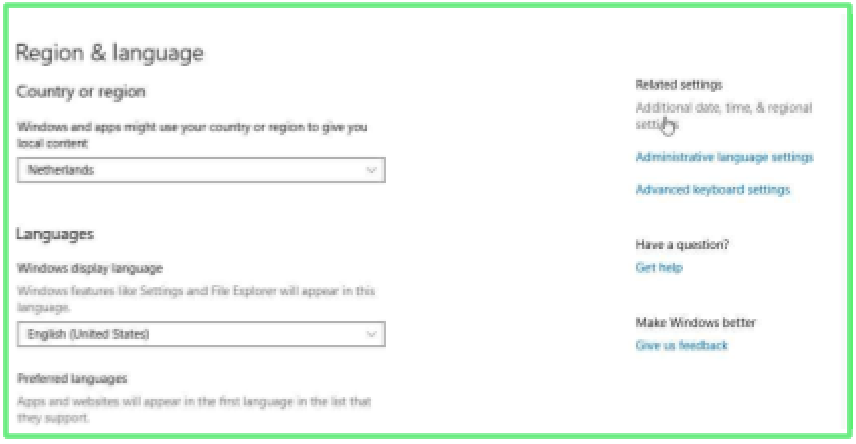
The settings of your computer must be in the right notation. The settings of your computer are so that the dot (.) is the decimal symbol, while the comma (,) is the digit grouping symbol (i.e. a million is 1,000,000.00). For most computers this is already the default, otherwise, follow the steps below. You might skip this step but if you get really strange results later, please try this step first to solve this issue.

Depending on your Windows version the following steps might look different at your computer. The figures below are for Windows 10. To check if the settings are correct:

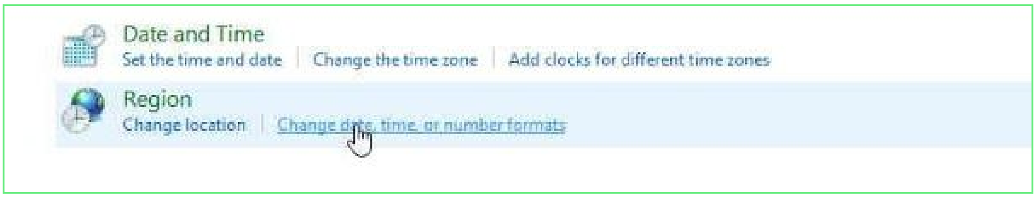
1) Go to the settings of your computer


Best match
Settings
App
Apps
Windows Language Settings
Search the web
sett - See web results
Settings (7+)
People (1+)
settings

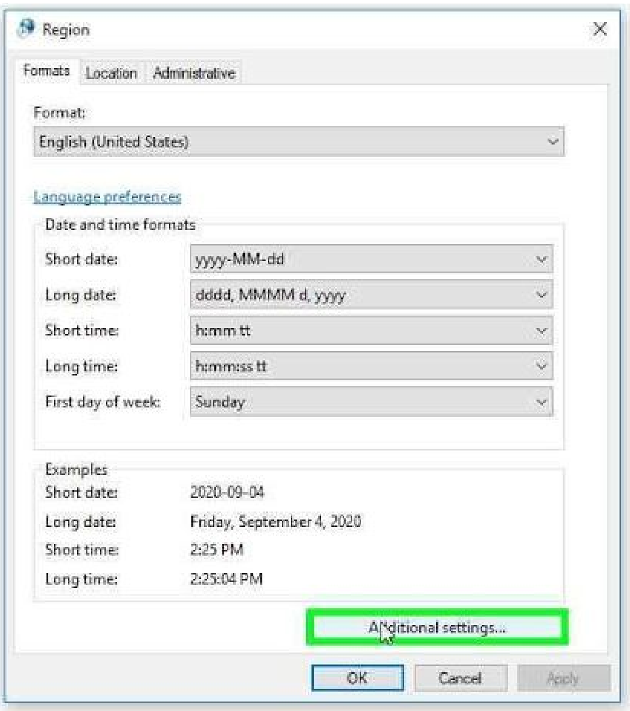
2) Go to Region and language --> additional date, time & regional settings



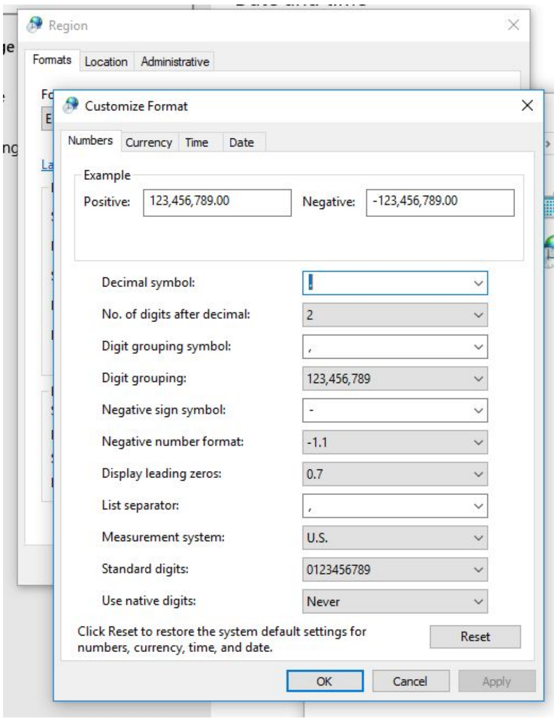
3) Click on “Change date, time, or number formats”



4) Click on “additional settings”



5) Check if your settings are according to the figure below. If not, adjust.



decimal symbol = .

digit grouping symbol = ,

negative sign symbol = -

List separator = ,

Measurement system = U.S.

STEP 8: Prepare your PCLake project

For the models, a simple interface and model code can be found in an EXCEL spreadsheet located at:

PCLake+ (or shallow and deep lakes)

.../Licence_agreement/I_accept/PCModel1350/PCModel/3.00/Models/PCLake-plus/2.13.16/PL613162.xls

PCLake (only shallow lakes, this version is not updated anymore)

.../Licence_agreement/I_accept/PCModel1350/PCModel/3.00/Models/PCLake/2.13.16/PL613162PLUS.xls

PCDitch (for linear waters)

.../Licence_agreement/I_accept/PCModel1350/PCModel/3.00/Models/PCDitch/2.13.16/PD213162.xls

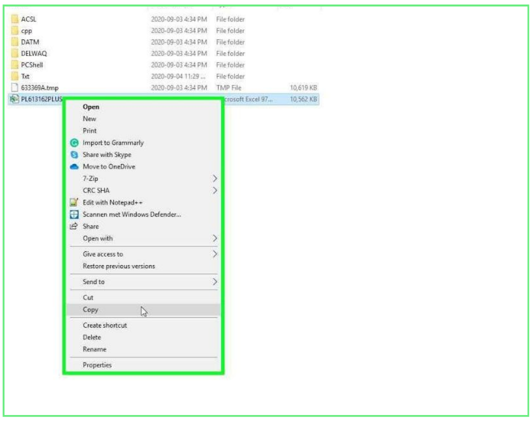
From now on we only focus on PCLake+. The other models work similarly

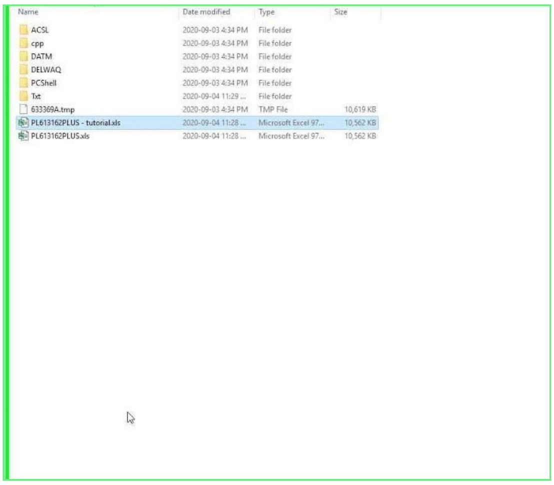
1) Go to the path:

   ...\PCModel-master\PCModel-master\Licence_agreement\I_accept\PCModel1350\PCModel\3.00\Models\PCLake+

2) Make a shortcut of the folder called “6.13.16”

3) Paste the shortcut at your desktop (or other easy location)
4) For safety reasons make a copy of the file “PL613162PLUS.xls”. You do this each time you start a new project or even within the project to make sure that your changes will not be lost by new changes.
5) The name of the new file should be something like PL613162PLUS<your addition>.xls. Never forget to keep “PL613162PLUS'' at the beginning of the name. Otherwise, you will get run errors. (in the example below the name became PL613162PLUS - tuturial.xls). Make sure that the xls files are stored at the same location in the directory tree as the original Xls-file.



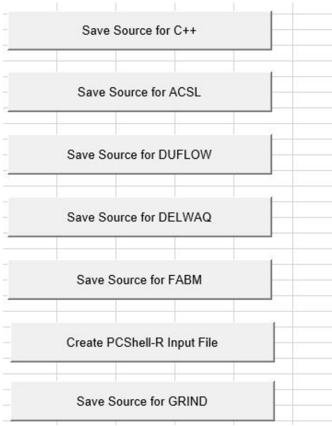


STEP 9: Database Approach Towards Modelling (DATM)

Here you read some background on how PCLake is set up.

The Database Approach to Modelling (DATM) is applied to run PCLake. An explanation of this approach is presented in an open-access article in Environmental Modelling & Software: http://dx.doi.org/10.1016/j.envsoft.2014.04.004

The idea of DATM is that the storage of ecological knowledge is independent of the language and platform in which the model will be run. In this case, Microsoft Excel is used as a database program to store the information of PCLake, PCLake+ and PCDitch. To create a working version of PCLake or any other of the three models in a specific software program, the information in the database is translated and augmented with the language and platform specifics. This process is automated so that a new instance of the model can be created each time the database is updated. You can translate your files by pressing one of the buttons below:



Using DATM it has become possible to run PCLake, PCLake+ and PCDitch in OSIRIS, GRIND for MATLAB, ACSL, R, DUFLOW and DELWAQ. Translated files are stored in the folder "6.13.16/<your chosen language>". Only the first 'default' set of each states, parameters, initstates and derivs field is used in the translated PCLake. If you want to translate your customized code, either copy the changed fields into the first column, or change the translator in the associated VBA macro. The most used options are PCLake in R, PCLake in GRIND for MATLAB and PCLake in SOBEK II or III. We assume you are an expert in these languages once you choose this option. Therefore, the exact procedure for using this option is not part of this tutorial.

The essential information needed to run PCLake in any given simulation platform is stored in five tables/sheets:

a) Simulation information, to control the simulation '*control*',
b) Values of the state variables at the start of a simulation(t=0), cf. initial states are in '*states*',
c) Parameter values '*parameters*',
d) Dynamical equations '*derivatives*' (calculation of auxiliaries and derivatives).
e) Equations that are only used once at the start of the simulation '*initialstates*',

These latter initial equations are used, for example, to ensure that the initial values of the state variables obey certain basic biological rules (e.g. stoichiometric constraints) in case initial values are provided only for dry-weight but not for N and P. Also, these equations set the initial composition of the sediment (this is not very important to understand for a beginner). For PCLake+ also the initial states equations can be used to transfer initial states given in concentrations into mass values.

STEP 10: Open PCLake

Now it is time to start with your first run.

Open the xls-file you just copied in the folder "6.13.16" in the previous STEP 8. If asked: answer 'yes' to 'Enable macros?' or 'Enable content?'.

The first thing you see is a worksheet with graphs, buttons and a table with some numbers. This is the *'control'* sheet. We will figure out later how we are going to use this control sheet, but let us first verify that also the other sheets are present. Check the tabs with *'states'*, *'parameters'*, *'initialstates'* and *'derivatives'*.

Open the *'states'* sheet, which presents a list of all the state variables of PCLake together with the values of the state variables at (t=0). Together these values give a static representation of the ecosystem at the start of the simulation. If you look at the sheet you notice that in the middle there are four columns with values (columns I-L). For PCLake to run we would need only one list with values as obviously, each state can have only one value at t=0. However, with four lists we can easily switch between simulation setups, for example between clear and turbid initial conditions. By default, the first list with initial state values represents a typical turbid lake that is representative for many turbid lakes in the Netherlands (column I), and the second list represents a clear lake with macrophytes (column J). The two remaining lists are duplicates of the turbid lake.

Now look next at the other columns that are present in the *'states'* sheet. A column named *sStateName* holds the names of the state variables, the column with the name *sStateUnit* holds the units and column with the name *sDescription* holds a short description of each state variable. The other columns provide meta-information and are not important at this moment.

Now check the '*parameter'* sheet and notice that this sheet also comprises four separate lists with parameter values (column J-M) and columns with names (column F), units (G) and a description of each parameter (T). There are columns which hold additional information but these are not important at this point.

Finally check the '*initialstates*' and '*derivatives'* sheets. In the '*derivatives'* sheet the auxiliary equations and differential equations are stored along with three duplicates. The user may want to modify certain equations which would imply that four different versions of PCLake can be presented in this sheet side by side which allows for easy comparison.

Tip: NEVER change the order of the equations e.g. by changing the alphabetic order!

At this stage, we have obtained an idea of how DATM works and witnessed that all the information that is necessary to produce a working model version of PCLake is present. The next thing is to get going. From here there are different routes. In the '*control*' sheet at the lower left, we can find different buttons which say: "Save Source for…". By pressing these buttons we can produce working code for, for example, Matlab (GRIND) and R. However, one needs some programming experience to run PCLake in these platforms. Fortunately, the Excel file we just explored is coupled to the OSIRIS framework, which allows us to run PCLake from Excel! In the next chapter, we will learn how to operate the PCLake and run simulations.

PART 2B: RUN PCLAKE

STEP 11: Your first PCLake run

The DATM implementation of PCLake in Excel is directly linked to the OSIRIS modelling framework. This enables a user to run PCLake directly from Excel. Hence, besides being a database our Excel file is also a user interface. What happens is that the information in the Excel sheet is translated into text files (.txt) using VBA macros in Excel. These .txt files are then automatically imported by a piece of software that is produced by OSIRIS (a C++ executable). This software solves the equations and produces model output. The output is produced as txt files which in turn are automatically imported back into Excel where they can be accessed and processed by the model user. Now let's get started.

The first thing we want to do is to perform a simple simulation with PCLake whereby the dynamics of a lake are modelled as a function of time. To do so, we need to tell the program several things such as which equations should be used, which parameter values should be used as well as which initial values for the state variables should be used. Also, we need to tell the program the length of the simulation and specify what kind of results we want to see. Basically what we want to do is to control the simulation.

Control the run

Open the '*control*' sheet. The first things you probably notice are the fifteen graphs showing model output. On the left side of these graphs, there are several buttons and there is a table. In this table, you can specify your run (see figure). There are five columns in the control table; four of them saying Set 0, Set 1, Set 2, Set 3, respectively. These four sets stand for four separate simulations that can be run simultaneously, side by side. Each set, or simulation, needs to be specified. In the first column, which says CONTROL, the control options are listed. The meaning of each control option is explained and presented when the mouse indicator is placed on the red triangles.

Take for example the control option *dReady*. Here we can specify the duration of the simulation we want to run. The unit is years. Now if we look at the setup of the first simulation we see that the default value of *dReady* is 25 years. Hence these latter simulations are currently not used.

	Set0	Set1	Set2	Set3
iRuniD	0	1	2	3
iIntType	3	3	3	3
iStateSet	0	1	0	0
iParamSet	0	0	2	3
iInitAuxilSe	0	0	0	0
iAuxilset	0	0	0	0
iCalibType	0	0	0	0
iCalibMaxI	50	50	50	50
dReady	25	25	25	25
dIntStep	0.001	0.001	0.001	0.001
dIntAccura	0.0001	0.0001	0.0001	0.0001
dRepStart	0	0	0	0
dRepStep	1	1	1	1
dAvgStep	1	1	1	1
dAvgStart	24	24	24	24
dAvgStartV	150	150	150	150
dAvgEndW	210	210	210	210
dSensitivit	0.01	0.01	0.01	0.01
dCalibScal	0.05	0.05	0.05	0.05
dCalibAccl	0.001	0.001	0.001	0.001
iGccComp	c:\rtools\gcc-4.6.3			
iFilesDATM	DATM\PCLakePLUS_20200828\			

The next thing we want to do is to specify which parameters, initial state variables and differential equations we want to use during the simulation run. Let us focus on the simulations *Set 0* and *Set 1*. We then look at the control option *iParamSet*. Now, remember from the previous step that in the *'parameter'* worksheet four different lists with parameter values were present. With the control option *iParamSet* we can define which of the four parameter lists we want to use during this simulation. We can enter the numbers 0, 1, 2 or 3, which correspond with the columns J, K, L, M respectively, in the *'parameter'* sheet. By default, the '0' is filled in, and the simulations thus make use of the first list with parameter values (column J). But we could just as easily select one of the other parameter lists by filling in a 1 or a 3, which would then allow us to compare different simulations using different parameter lists.

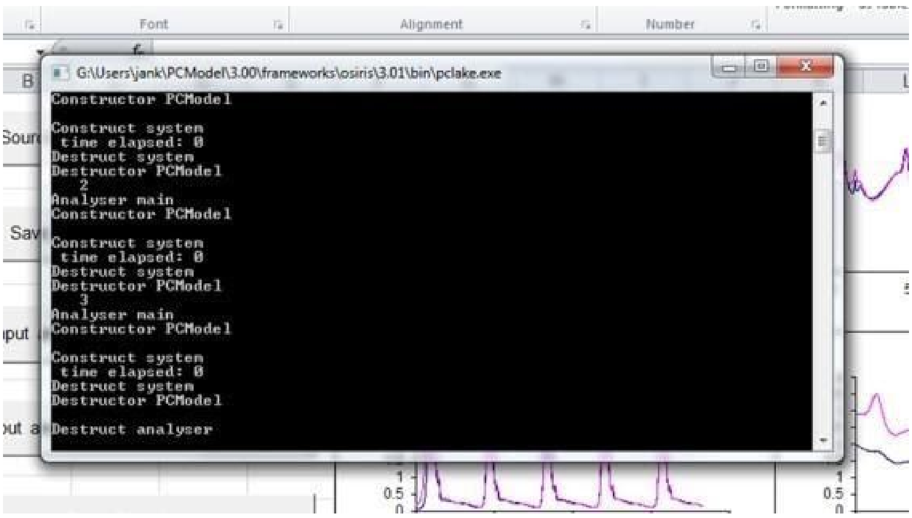
The same we can do for the control option *'iStateset'* which enables us to select one of the four lists with initial values of the state variables presented in the '*states*' worksheet. Again the numbers 0, 1, 2 or 3 correspond to the four columns with initial values that are present in the '*states*' worksheet. If you look at our simulations, *Set 0* and *Set 1*, you will note that they point to different lists with initial state values. *Set 0* makes use of the first column (0) while *Set 1* includes the second column (1). From the previous chapter, you may remember that the first column represents a turbid lake, while the second column represents a clear lake.

Question: By looking at the default output graphs, can you identify which of the two lines belongs to the simulated lake starting with a clear state and which one belongs to the turbid state? What happens with the clear lake using the current parameter values, will it stay clear or does it become turbid?

One of the control options is *dRepStep.* Here we can define the frequency of our output. If we fill in a 1 (default) the program gives us output data for every day. If we fill in a 2 the program produces output for the second day. Important to note here is that the program will only create 2000 output records for each scenario run. So if we fill in a 1, the program is only to show us 5.5 years. However, if we ask for output every 2 days, we can get output for 11 years, and if we ask output for once a week, we can simulate 38 years.

Start the simulation

After defining a simulation we can perform the simulation run. To do so, press the button "Save Input and Run Model". A black screen appears, and after a short moment (depending on the duration of the simulation, many years means longer calculation time) it automatically disappears.



If the run ends with “destruct analyser”, you have done a good job. Yet you may also see errors in this black screen. In that case, it means you have to check your input.

An important next step is to press the button “Open Output”. Although our software program has already produced output as txt files, we need to import the information into Excel by pressing this button.

Now if you have not changed anything in the control settings or in the parameter lists, nothing will change after clicking on the “Open Output” button: the simulation setup stays the same and hence the output is replaced yet nothing changed.

Select output

In the ‘Control’ sheet we find 15 graphs with the output of important variables, such as the concentration PO4 in the water column in the epilimnion layer (_*sPO4WEpi*_). PCLake in Excel does not automatically produce output for every variable. The output of state variables is *only* produced for those variables that were *marked with a ‘1’*, in the *‘iReportState’* column in the *‘states’* sheet.

Go to the ‘*states*’ sheet, and find the column (C). Now check that most state variables are marked with a 0, but that some are marked with a 1. You can make changes in this column if you also want output for variables that are currently not selected (change a 0 to a 1).

Importantly, it is also possible to get output for auxiliary variables. Basically, all values that you want as output in concentrations can be found in the derivatives sheet. Also, chlorophyll-*a and Secchi depth* can be chosen as output from the derivatives sheet. Besides you can ask for the output of processes such as when you want to know how much vegetation is consumed by birds. For all these cases, go to the ‘*derivatives*’ sheet where you find a column that is called ‘iReport’. Also here it is possible to mark equations with a 1 if their output is desired.

Finally, the parameters can also be plotted, which is helpful if forcing time series are used, for example, the external nutrient loading that may change over time. Also in the ‘parameter’ worksheet, there is the ‘iReport’ column, where parameters can be marked with a 1.

For the first 18 variables output data is presented in graphical form on the control sheet. However, the output for all selected variables appears in tabular form in the sheet ‘*sysrep’*. With this primary data, you can make your own graph in Excel, or use Excel to calculate some basic statistics.

Tip: This data in sysrep can also be found in the folder “.../6.13.16/txt” in the files called systrep0, sysrep1, sysrep2 and sysrep3 (if you run all the four sets). These files are not limited to 2000 entries which is the case for systrep in the excel file. This tip is relevant for longer runs not part of this tutorial though.

Question: Choose a state variable, parameter or equation from the auxiliary that is currently at 0 at set it to 1. Run the model and check if it appears in the figures given in the control sheet.

Adding your data

To become a PCLake expert, it is important that you play a lot with the model: make practice hours. Just change things and see what happens. This can best be done by changing one thing at a time and see what is the effect of that. Note that the exercises below work best if you have not changed anything yet in the Excel file. If you want to be sure of this you can make a new copy of the original excel file downloaded from github and start with this fresh copy.

Go to "*control*" to change the 'dReady' to the following values.

dReady	25	25	0	0

Also set dRepStep to the following values

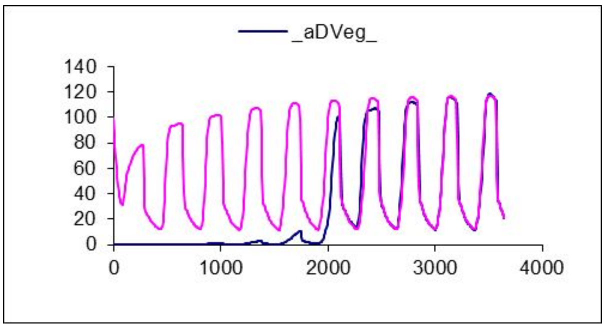
dRepStep	7	7	1	1

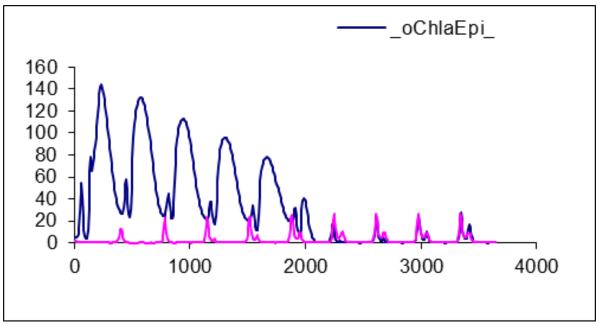
Question: What do these values for dReady and dRepStep mean?

Setting the dReady to 0 means that the run is not used. This saves time while running the model.

Change the depth of the lake. Go to the '*parameters'* sheet and look for _cDepthWInit0_. The default value is 2 meters. Now change the depth of the lake from 2 meters to 1.5 meters (in both clear and turbid) and run the model. Don't forget to open output after the calculation is finished.

Question: Are the results for the pink and blue line different for the last years?



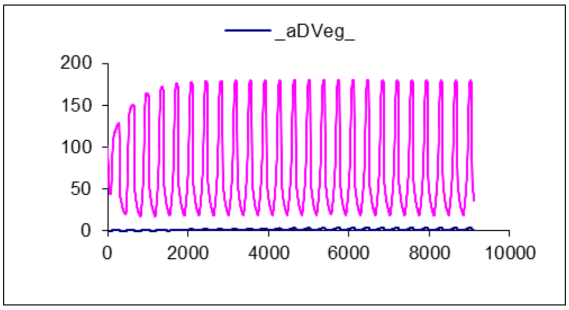


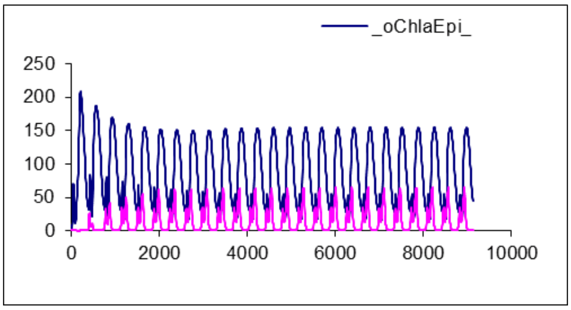
Probably your answer to the question is that both lines show the same output for the last year. With the current parameter settings, the 1.5-meter deep lake turns clear, even if the initial state variables are turbid.

Now we are going to increase the nutrient loading (keep the depth at 1.5 m). In the 'parameter' sheet we can manually modify the parameter values. Search for the _mPLoad_ and read the description and the unit. The default value is 0.001 g/m2/day. Now change this to 0.0025 g/m2/day. You can do this in the first column with parameter values, or in one of the three other columns which are replicates. It does not matter in which one you change the parameter values as long as you point to the corresponding column in the control sheet. For this exercise, it is easiest to change only the first columns, as simulations Set 0 and Set 1 are currently pointing to this column. Then run the model again and open output.

Question: Now that do you see?

If everything worked out, the results now show that a lake starting with clear initial values stays clear with macrophytes, while a lake starting with turbid initial values stays turbid with high chlorophyll-a concentrations. It is important to note that in both simulations the parameter values are identical! The only different thing is the initial values (which also represent the history of the lake). This is a typical feature of alternative stable states. For specific environmental conditions, both the clear-water state and the turbid water state can exist as alternative stable states and the prevalent state depends on the foregoing conditions.





STEP 12: Adding time series

In the '*parameter*' sheet the values of parameters are listed. These are fixed values. Sometimes however we like to use forcing time series, for example for the nutrient loading, irradiation, water inflow, evaporation etc. Thus far we have seen the _mPLoad_ which was a parameter with constant values. However, the actual nutrient loading is likely to change over time, and if the data is available it would be beneficial to take this information into account.

Check the worksheet 'mPLoad'. This sheet is an example sheet for a time series of P loading.

	A	B	C
1	dTime	dValue	-1
2	0	0.0006	-1
3	89	0.0006	-1
4	90	0.01	-1
5	200	0.01	-1
6	201	0.0006	-1

The data in this sheet is presented as follows: column A = time, column B = values. Missing data are linearly interpolated. The first line should contain the words "dTime" and "dValue" as column headers. Column C should contain "-1" on all lines; the next columns may be used for comment if desired. The table is also closed by a "-1" on the next line. It is easy to copy this sheet and change the name into for example _mPLoad_, to create an input file for a time series of N loading. Afterwards, change the content of the sheet by copying the N loading data into the sheet. Do not forget the -1 signs.

The model reads the data from a text file and not from the spreadsheet itself; these text files are written when the button 'Save input and run model' is pressed. However, what is important is that we need to notify PCLake that it needs to use these txt files with time-series data. To do so, two important steps need to be taken.

1) *Set a switch/flag.* For the most used parameter forcing's (temp, wind, light, inflow, outflow, evaporation, nutrient loading, loading of organic/inorganic matter), there is a switch/flag that needs to be set in the parameter list, defining whether the model needs to use a constant value or measured time series. For example, _ReadPLoad_: the switch is set to '0' if you want to use a constant P loading or to '1' if you want to use the measured P loading. In other words, if _ReadPLoad_ = 0, then PCLake uses the constant value defined by the parameter _cPLoad_ (c is for 'constant'). If _ReadPLoad_ = 1, then PCLake uses the value(s) defined by the parameter _mPLoad_ (m is for 'measured').
2) *Point to the right file.* By default, the parameter _mPLoad_ has the value 0. But of course, we do not want the P loading to be 0, but we want the time series data to be used. What we need to do is replace the 0 with the following statement: txt/mPLoad.txt

This statement tells the model that it needs to use the txt file with time-series data. Be careful that the name in the statement is identical to the name of the sheet: _mPLoad_.

Depending on the available data, we could include worksheets with time series for temperature, wind, light, inflow, outflow, evaporation, nutrient loading, organic/inorganic matter input, set all the corresponding switches in the '*parameter*' sheet, and present all the right statements to refer to the sheets. Remember to make sure that when setting up the simulation run in the control sheet you point to the right parameter list where you made all the above changes.

Also, make sure that the time series you use are long enough. For example, if you run for 25 years, the time-series should also be 25 years.

Tip: The keys Ctrl-PageUp and Ctrl-PageDown offer convenient ways to navigate between sheets.

STEP 13: Your first lake

Now take a fresh PCLake/Excel file and set up your own case study. Take it to step by step: one makes mistakes when being too hasty. To change a parameter and run the model, check if the output shows a pattern as expected. For example, if you change the depth, look in the output if the state variable for depth also changed. If you change your water input, check if the auxiliary for the water input takes up this new value when you rerun the model.

Also, make regular backups of your file, if you do run into troubles (a crash cannot be ruled out), you will have previous versions saved. Take the default values if you do not have measured values. The world is complex, so don't expect wonders from a model like PCLake, but most certainly it will help you to grasp a bit of the complexity and improve your understanding.

Tip: there is a lot to learn by examining the existing reports and publications of PCLake studies.

Question: To get experience with your own case, you can start with changing the parameters you found in PART 1 for a hypothetical case. Do this step by step (so not changing all the parameters in once). This helps you to trace back the reason for possible errors and also to understand why the model provides the output you get.

Check each time whether you understand why a certain change in parameters gives a certain result in your output. <u>Store your file once you are finished. You can use a copy of this file in STEP 16</u>

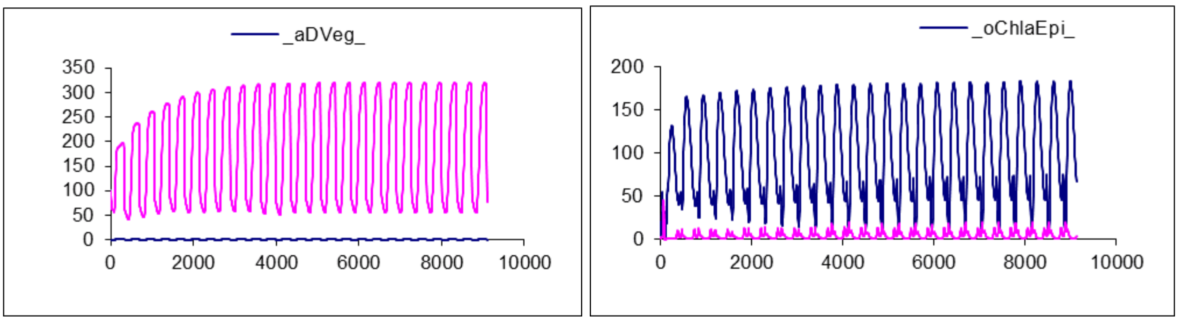
Parameters should look like below in this question on PART 1:

Bifurcation	sName	sUnit	sMinValue	sMaxValue	sDefault0
0	_ReadQIn_	0	0	0	1
0	_ReadQEv_	0	0	0	1
0	_mTempEpi_	_oC_	0	0	txt/mTempEpi.txt
0	_mQIn_	_mm_*_d_^-1	0	0	17.6624514
0	_mQEv_	_mm_*_d_^-1	0	0	1.28407589
0	_mPLoad_	_gP_*_m_^-2*_d_^-1	0.0001	0.01	0.004
0	_cFetch_	_m_	800	1200	270
0	_fLutum_	0	0	0	0.4
0	_fFeDIM_	_gFe_*_gDW_^-1	0	0	0.04
0	_fAlDIM_	_gAl_*_gDW_^-1	0	0	0.04
0	_fDTotS0_	_gDW_*_gDW_^-1	0	0	0.2
0	_cNPLoadMeas_	_gN_*_gP_^-1	0	0	13
0	_fDOrgSoil_	0	0	0	0.165
0	_cLAT_	_deg_	-90	90	41.1

The sheet mTempEpi looks like:

	A	B	C	D
1	dTime	dValue	-1	
2	0	0.1	-1	
3	30	0.2	-1	
4	60	2.8881	-1	
5	90	8.8821	-1	
6	120	15.213	-1	
7	150	18.714	-1	
8	180	24.596	-1	
9	210	25.13	-1	
10	240	16.099	-1	
11	270	11.299	-1	
12	300	7.5297	-1	
13	-1			
14				
15				
16				
17				
18				

The results should look like:



Question: what do the results mean? You can check what happens if you change the PLoad.

If you want later to simulate your own lake, you have to prepare your data according to PART 1 of this tutorial.

Once PCLake is fully set up, many analyses can be performed, depending on the goal of the modelling study, such as validation, sensitivity analysis, bifurcation analysis, uncertainty analysis, calibration, validation, scenario analysis etc.

Experience has taught us that in the vast majority of cases, the process of gathering all the necessary data and setting up the model to simulate a specific lake will already be very instructive and informative. In several cases, ecosystem managers were able to find a solution for their water quality problem even before simulations were run, just by taking the model as a consistent framework where all their knowledge was gathered and examined. Working with PCLake, PCLake+ and PCDitch provide basic and valuable insight into how aquatic ecosystems function. Although the ultimate goal of a modelling study may be to make accurate predictions of future scenarios, the biggest goal will always be to obtain understanding. Only when there is understanding we can continue with prediction.

STEP 14: Final note on running transient output

The simulations that you have done above are called transient-simulations. Transient simulations show output over time. Please note that:

1. Transient output depends on their initial values. To estimate the initial values of your lake you can run PCLake first with all settings for the first year, but repeating that for 50 years. If the 50^{th} year is similar to the 49^{th} year for most/actually all of the state variables you can continue, otherwise you have to run your model a bit longer. You can use the values of 31^{st} of December of the last year of your run as initial values. To do so you can either copy the values to the states sheet (check the units!!), or, more simply you can start the rest of your time-series after the 50^{th} year/ or the last year you found them to be equal. So your real run then starts at year 50.
2. The transient output is used to
   a. Validate the model for your own case
   b. See the response time of specific measures
   c. See when the model comes to equilibrium (important for part 3)
3. Before reporting your output in a report or paper. Always check the values of_aDError_, _aNError_, _aPError_, _aSiError_, _aO2Error_. If they are too high (> 1e-6) it means you have a mass balance error. First, solve this mass balance error before you proceed.

PART 3: RUN BIFURCATION

STEP 15: Run a bifurcation for the default settings

PCLake is particularly useful to determine the critical values of external stress that lead to sudden shifts in qualitative model behaviour (i.e. bifurcation, catastrophic shift, critical transition). To test the effect of a specific external variable (e.g. nutrient loading), the model is evaluated for a range of parameter values (e.g. increasing nutrient loading) and for different initial conditions (clear vs. turbid). The equilibrium output values (after approx. 20 years) of certain dependent variables (e.g. summer averaged chlorophyll-a, can be used to evaluate the state of the system. When a small change of the parameter value leads to a large change of the dependent variable, we have located our 'tipping point'. See the example in the figure from Janse et al 2010.

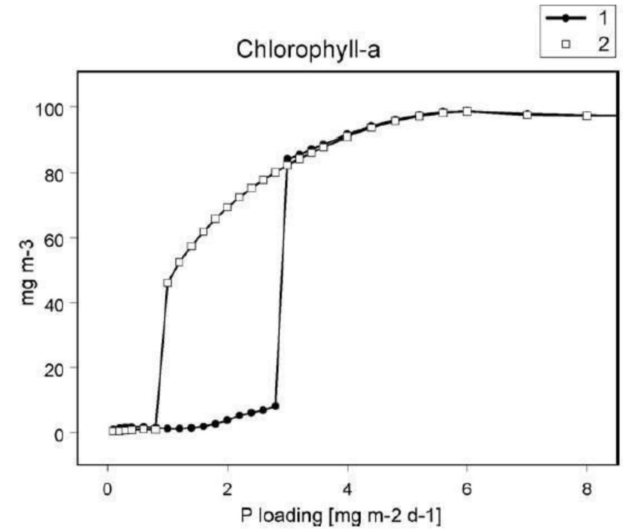


Fig. 4. Chlorophyll-*a* as a function of P loading, for two different initial states: 1 = clear state; 2 = turbid state.

The DATM implementation of PCLake in Excel has a feature built-in that is going to help us to perform this simulation exercise. This application starts if we press the button 'Save Input and Perform Bifurcation Analysis'. But before we do so, we first have to take several steps to set up the simulation experiment.

Start with a fresh workbook, with all states and parameters set to default values.

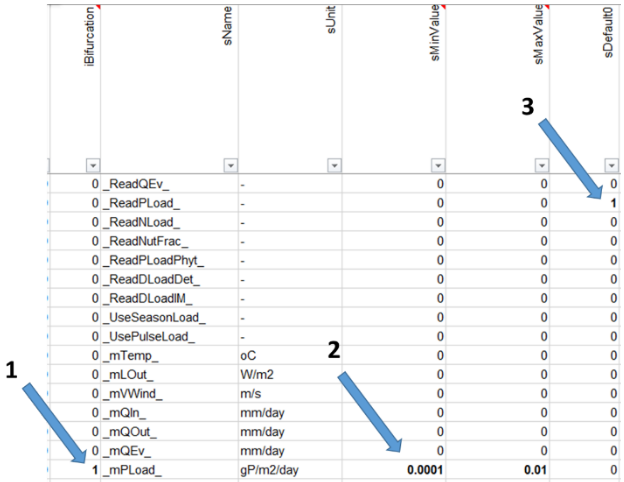
1) Select the clear and turbid start conditions: the bifurcation application uses *Set 0* and *Set 1* to run its simulations. Go to the '*control*' sheet and check for *Set 0* if the iStateSet is pointing to the turbid water state (0) and for *Set* 1 that the iStateSet is pointing to the clear water state (1).
2) Define how long the simulations need to run. We want to evaluate the system in equilibrium. Normally 30 years is sufficient to reach equilibrium, but if you have deeper lakes it might take longer. Experience showed that some lakes even need 100 years. You can use the transient calculations to see if the model goes to equilibrium. We here make sure that *dReady* of both *Set 0* and *Set 1* at 30 years.
3) Define at which stage in the simulation the averages should be calculated. In the last year of the simulation, we will assess the state of the system. The state is assessed by taking year or summer averages.

The control option *dAvgStart* defines the year when averages should be calculated. For both *Set 0* and *Set 1* we change the *dAvgStart* from 9 years to 29 years, which is *dReady* minus one.

4) Define which part of the year should be averaged. Often we are interested in the summer average of the chlorophyll-a concentration in the final year of the simulation to characterize the state of the system. We can use the control options *dAvgStartWithinYear* and *dAvgSEndWithinYear*. For both Set 0 and Set 1, check that these are on 150 and 210 respectively.

'control' sheet	Set0	Set1
iRuniD	0	1
iIntType	3	3
iStateSet	0	1
dReady	30	30
dAvgStart	29	29
dAvgStartWithinYear	150	150
dAvgEndWithinYear	210	210

5) Define the independent bifurcation parameter (x-axis). The independent bifurcation parameter can be defined in the *parameters* sheet, by marking the parameter with a '1' in the '*iBifurcation*' column (column E, arrow 1 in the figure below). Also, the range should be defined over which the bifurcation parameter is varied (i.e. the minimum and maximum value) in the column '*sMinValue*' and '*sMaxValue*' (arrow 2 in the figure below). As PCLake is mostly used as a eutrophication model to calculate the critical nutrient loadings, the nutrient loading is typically used as the independent control parameter. Normally we use the *_mPLoad_* and not the *_cPLoad_*. One hidden feature of the *_mPLoad_* is that the N loading will be automatically coupled to the P loading with a ratio of 10, unless the _mNLoad_ is also specified. For this bifurcation analysis, it implies that both the N and P loading change simultaneously, making that the system remains P-limited during the analysis. If only the P loading would increase while the N loading would stay the same then the system would switch from a strongly P-limited to a strongly N-limited system during the analysis. We are rather interested in the combined eutrophication effect. To communicate to the computer that the *mPLoad* is used instead of the *cPLoad*, remember to mark the *ReadPLoad* with a 1 in the parameter column (arrow 3 in the figure). For the '*sMaxValue*' (arrow 2 below) choose 0.005.



iBifurcation	sName	sUnit	sMinValue	sMaxValue	sDefault0
0	_ReadQEv_	-	0	0	0
0	_ReadPLoad_	-	0	0	1
0	_ReadNLoad_	-	0	0	0
0	_ReadNutFrac_	-	0	0	0
0	_ReadPLoadPhyt_	-	0	0	0
0	_ReadDLoadDet_	-	0	0	0
0	_ReadDLoadIM_	-	0	0	0
0	_UseSeasonLoad_	-	0	0	0
0	_UsePulseLoad_	-	0	0	0
0	_mTemp_	oC	0	0	0
0	_mLOut_	W/m2	0	0	0
0	_mVWind_	m/s	0	0	0
0	_mQIn_	mm/day	0	0	0
0	_mQOut_	mm/day	0	0	0
0	_mQEv_	mm/day	0	0	0
1	_mPLoad_	gP/m2/day	0.0001	0.01	0

6) <u>Select the dependent output variable to evaluate the state of the system</u>. This can be done in the '*derivatives*' sheet, also by marking the variable with a '1' in the '*iBifurcation*' column (E). Often chlorophyll-a is used (_*oChlaEpi*_). But the total plant biomass would also be a good option to evaluate the state of the lake (_*aDVeg*_).
7) <u>We have now specified everything we need.</u> Thus press ***'Save input and perform bifurcation analysis'*** button in the 'control' sheet. Then '***Open Output'***, and find the results in the '*bifurrep*' sheet.

Question: How does the bifurcation plot look like? What are the critical nutrient loads?

Question: Perform several bifurcation analyses to analyze the importance of one of the lake characteristics, for example by making the lake deeper or by increasing the fetch. Explain the results from the feedback diagram; which feedbacks are involved? What are the implications for lake ecosystem management?

Question: Perform several bifurcation analyses to analyze how temperature influences the resilience of the lake. Explain the results from the feedback diagram; which feedbacks are involved? What are the implications for lake ecosystem management? What kind of measures would be effective?

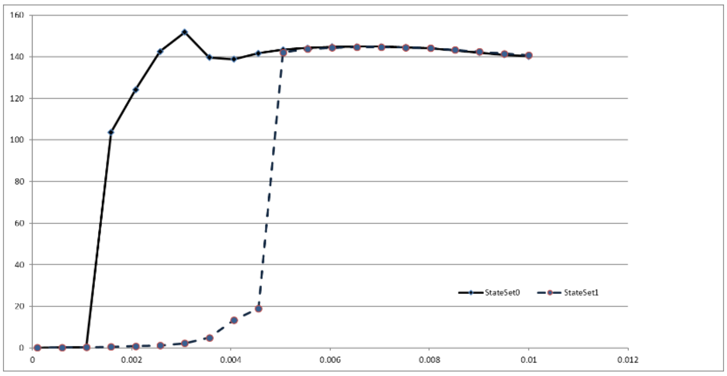
STEP 16: Bifurcation for your first lake

If you have stored your file in STEP 13 you can do this exercise to make a bifurcation for your first hypothetical lake.

For a bifurcation analysis of the hypothetical lake, you have to take the same steps as in STEP 15 including the range of the PLoad, while keeping the parameter settings of STEP 13. The result will look similar like this:

Question: think of why this plot looks different than in step 15.

The example below shows that the line is not smooth yet (seen by the sudden peak at the dark solid line around 0.003 gP/m2/d). This means that the equilibrium was not reached yet. If you see this peak, you have to run your model longer to reach equilibrium. Do not forget to also change the year you want output for in "dAvgStart ".



STEP 17: Final note on running bifurcation analysis

To start a new bifurcation it is better to first estimate the new initials of your clear and turbid lake. To do this you have to take several steps.

1) Take the validated version of your PCLake xls file as a start (this can be as an example the version you had in step 13).
2) In the '*parameters*' sheet copy all parameter values of the validated lake to that of set 0 and set 1. This means that set 0 and set 1 are identical with respect to the parameter values that correspond to your specific lake.
3) Search for the _mPload_ and change the value for set 0 to a very high value (e.g. 0.01) and the value of set 1 to a very low value (e.g. 0.0001).

Bifurcat	sNa	sUn	sMinVal	sMaxVal	sDefaul	sSe	sSe	sSe
0	_mPLoad_	_gP_*_m_^-2*_d_^-1	0.0001	0.01	0.01	0.0001	0.001	0.001

4) Set your dReady in the '*control'* sheet to 50 for set 0 and set 1.

8	iCalibMaxI	50	50	50	50
9	dReady	50	50	0	0
0	dIntStep	0.001	0.001	0.001	0.001

5) Set your iStateSet and iParamSet and iStateSet in the '*control'* sheet to 0 and 1 respectively for set 0 and set 1.

3	iStateSet	0	1	0	0	1
4	iParamSet	0	1	2	3	

6) Set dRepStart to 45 to make sure you get the output of the last year

dRepStart	45	45	0	0

7) Set iReportState to 1 for all states in '*states*' sheet
8) Run the model and request output. Check if the 49th year looks the same as the 50th year. If not, run longer. If the last year looks the same as the year just before the last year, you can use the output of the last day (31st of December) as initial states. (next step)
9) Take the output from sysrep from all state variables of the last day and copy them to the state variables (e.g. set 3 and 4, to keep the original). Check if you copy the right units! It could be that you have to change some units!

10) You can check if you did it correctly by running the model with the new initial states while keeping the same parameter settings. If your model is immediately in equilibrium you have done the right steps.

PART 4: MODEL CHANGES AND COMPILATION

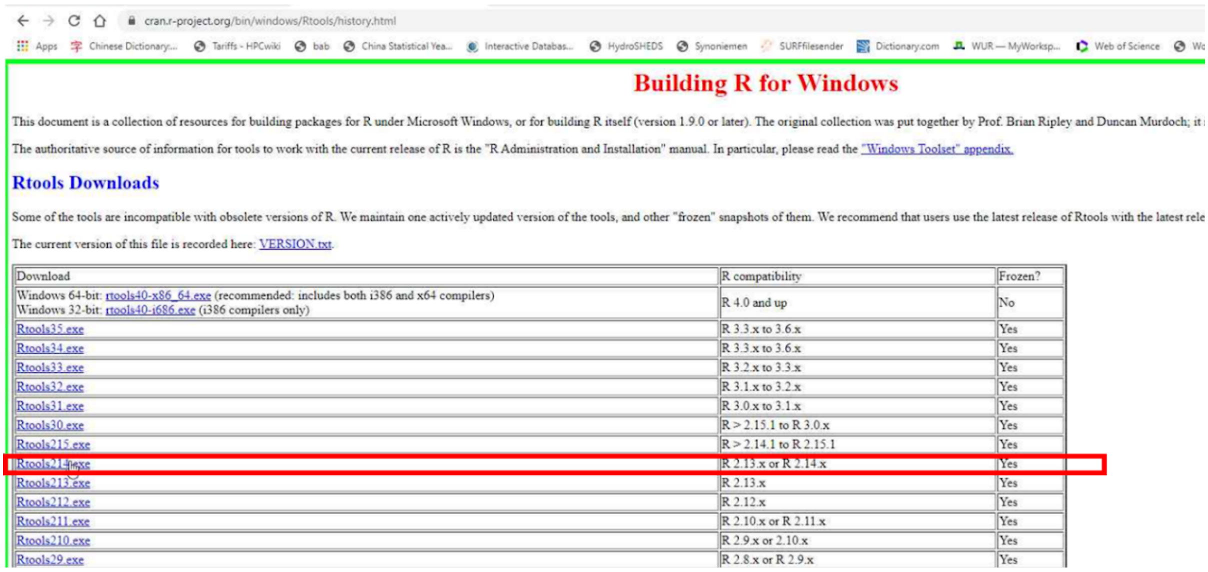
For most analyses, the current version of the PCLake model will be sufficient, even though not all processes are taken care of by the model. Yet, if one needs to change the equations of PCLake, it is possible. There are four columns with equations in the 'derivatives' sheet, which means that there are four replicate versions of PCLake available in the workbook, and this makes it easy to make some changes in one set of equations and compare the modified PCLake version with the original to see what the effect of the modification is.

You can make modifications in existing equations, but you can also add equations or parameters. This can simply be done by inserting a new row in the worksheet, however, a mistake is easily made. One type of mistake relates to communication with the OSIRIS framework. For example, do not leave any cell blank and make sure the row ends with a -1. Also, it is important to insert new equations in the correct order in the database. With this, we mean that each variable must be assigned a value before it is used in the assignment of another variable (in other words, it must first be used as a left-hand term before it is used as a right-hand term). As the second type of mistake can occur when equations violate the mass balance. For example when nutrients are removed from one state variable but not entered into a new state variable and thus disappear. This can be easily checked by looking at the error terms of the mass balance: _aDError_, _aPError_ and _aNError_, these should always be close to zero.

For any changes in the equations to take effect, PCLake first needs to be recompiled, which means that OSIRIS will produce a new executable. To do so, use the button 'Save source for C++ and Compile Model'. But first make sure that the free software package RTools is installed on the C drive. RTools includes a free GCC compiler which will be used by OSIRIS.

You can download RTools from https://cran.r-project.org/bin/windows/Rtools/history.html

Make sure you take the right version (see below)



cran.r-project.org/bin/windows/Rtools/history.html

Building R for Windows

This document is a collection of resources for building packages for R under Microsoft Windows, or for building R itself (version 1.9.0 or later). The original collection was put together by Prof. Brian Ripley and Duncan Murdoch; it

The authoritative source of information for tools to work with the current release of R is the "R Administration and Installation" manual. In particular, please read the "Windows Toolset" appendix.

Rtools Downloads

Some of the tools are incompatible with obsolete versions of R. We maintain one actively updated version of the tools, and other "frozen" snapshots of them. We recommend that users use the latest release of Rtools with the latest rele

The current version of this file is recorded here: VERSION.txt.

Download	R compatibility	Frozen?
Windows 64-bit: rtools40-x86_64.exe (recommended: includes both i386 and x64 compilers) Windows 32-bit: rtools40-i686.exe (i386 compilers only)	R 4.0 and up	No
Rtools35.exe	R 3.3.x to 3.6.x	Yes
Rtools34.exe	R 3.3.x to 3.6.x	Yes
Rtools33.exe	R 3.2.x to 3.3.x	Yes
Rtools32.exe	R 3.1.x to 3.2.x	Yes
Rtools31.exe	R 3.0.x to 3.1.x	Yes
Rtools30.exe	R > 2.15.1 to R 3.0.x	Yes
Rtools215.exe	R > 2.14.1 to R 2.15.1	Yes
Rtools214.exe	R 2.13.x or R 2.14.x	Yes
Rtools213.exe	R 2.13.x	Yes
Rtools212.exe	R 2.12.x	Yes
Rtools211.exe	R 2.10.x or R 2.11.x	Yes
Rtools210.exe	R 2.9.x or 2.10.x	Yes
Rtools29.exe	R 2.8.x or R 2.9.x	Yes

Tip: It is not allowed to point to an external forcing function (e.g. txt\mPLoad.txt) in the parameter list during compilation.

As an exercise you can try to add an equation for the trophic state index (TSI) as defined by Carlson (1977), where SD is the Secchi depth:

Carlson, R.E. 1977. A trophic state index for lakes. Limnology and Oceanography. 22:361-369.

TSI (SD) = 10 [6- Ln(SD)/ln(2)]

Make sure you put this equation after the line of the Sechi depth is defined in the auxiliaries sheet. After including this equation you can try to compile the model

PART 5: DEFINITION OF ALL WORKSHEETS IN THE EXCEL FILE

The Excel file of the DATM implementation of PCLake includes many worksheets. Some of them are important to the user while others are not and can be ignored. A short explanation of all the sheets can be found in the following overview:

1) states: initial conditions of the state variables (per run); setting output on/off
2) parameters: parameter values (per run); setting output on/off
3) states: initial conditions of the state variables (per run); setting output on/off
4) derivatives: setting output on/off for intermediate variables

5) sysrep: table of dynamic results
6) anarep: gives the average results
7) bifurrep: gives results of the bifurcation analysis.

8) Nom: explains the nomenclature of PCLake

9) analyser: for communication with the c++ executable. Definition of runs; setting time period for output averaging (automatically linked with the control sheet)
10) anasens: for communication with the c++ executable. Definition of runs; sensitivity analysis (use default settings)
11) anacalib: for communication with the c++ executable. Definition of runs; calibration (use default settings)
12) anabifur: for communication with the c++ executable. Definition of runs; bifurcation analysis (use default settings)

13) steprep, setrep, intrep, dmpall: not important
14) sensrep: gives results of the sensitivity analysis (for most users not important)
15) calibrep: gives results of the calibration exercise (for most users not important)
16) ACSL: holds the original ACSL code of PCLake programmed by Jan Janse.

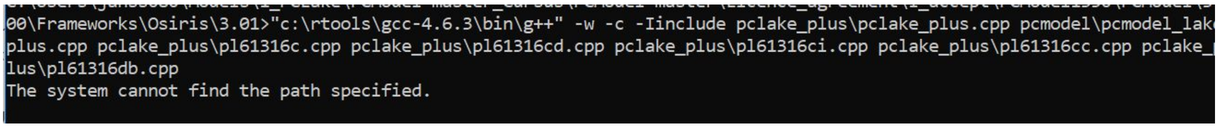
PART 6: FREQUENT ERRORS AND SOLUTIONS

Working with PCLake, PCLake+ and PCDitch may bring forward some technical questions. Unfortunately, there is no official help desk to ask about these questions. However, you can contact pcmodel@nioo.knaw.nl to ask for technical details. Below there is an overview of the most common technical problems:

COMPILING

Rtools not installed

- The system cannot find the path specified

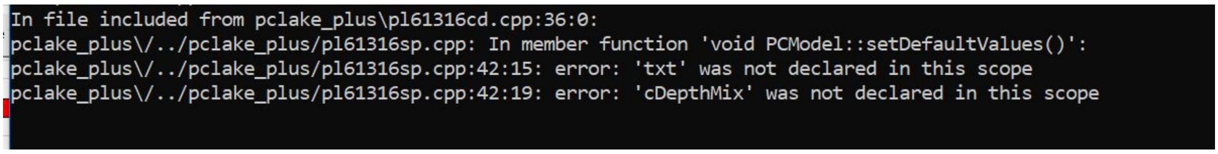


```
00\Frameworks\Osiris\3.01>"c:\rtools\gcc-4.6.3\bin\g++" -w -c -Iinclude pclake_plus\pclake_plus.cpp pcmodel\pcmodel_lak
plus.cpp pclake_plus\pl61316c.cpp pclake_plus\pl61316cd.cpp pclake_plus\pl61316ci.cpp pclake_plus\pl61316cc.cpp pclake_
lus\pl61316db.cpp
The system cannot find the path specified.
```

Rtools misses the right compiler. Go to PART 4 to identify how to install R-tools with the right compiler (gcc compiler)

- Error: ‘txt’ was not declared in this scope

Error: ‘parameter’ was not declared in this scope

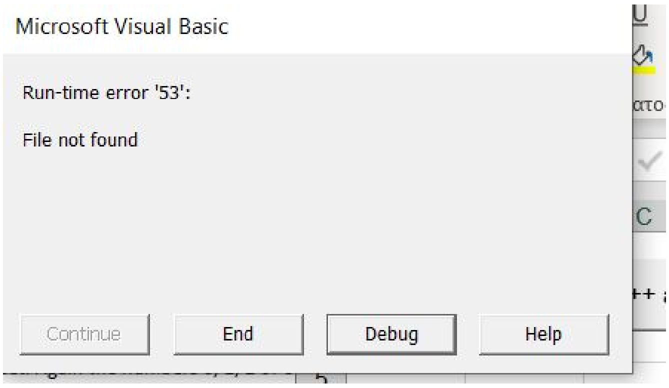


```
In file included from pclake_plus\pl61316cd.cpp:36:0:
pclake_plus\/../pclake_plus/pl61316sp.cpp: In member function 'void PCModel::setDefaultValues()':
pclake_plus\/../pclake_plus/pl61316sp.cpp:42:15: error: 'txt' was not declared in this scope
pclake_plus\/../pclake_plus/pl61316sp.cpp:42:19: error: 'cDepthMix' was not declared in this scope
```

The first error is because there is somewhere a time series defined (e.g. txt/mPload.txt) in the first column in the parameters sheet. To compile the model you always should make sure that the first column has only numbers (so no txt/<your_file.txt>)

RUN MODEL

- You get an that your file has not been found

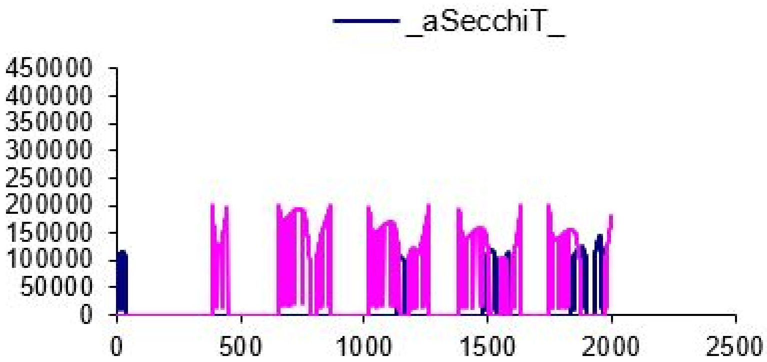


There might be several reasons that you did get this error:

1) *Your firewall deletes part of the PCLake model. If this is the case check your firewall software to solve this issue*

2) *When you download the PCModel you should always keep the same folder structure. Check if this is still the case. You can also download the whole pclake folder again and see if this error reappears, follow the steps in PART 4 to install R-tools and recompile your model again.*
3) *You pressed the button "Save Source for C++ and Compile Model" in an earlier stage without installing R-tools. To make sure this does not happen again,*
4) *Sometimes your previous run was not yet finished or you received an error that resulted in this error. You can try to close the xls without saving it (you might lose some data) and start it again. It might be that this error is solved now.*

- Your output gives very high numbers

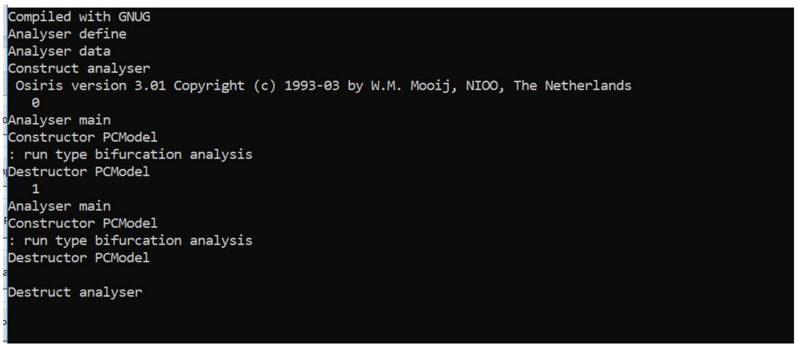


If your output provides very high numbers, this is because your computer settings are not correct. Please go back to step 7 in order to solve this issue.

- the -1 is missing somewhere
- the numbers in column A in one of the sheets are not counting up
- outside calculation range
- the depth is getting below zero

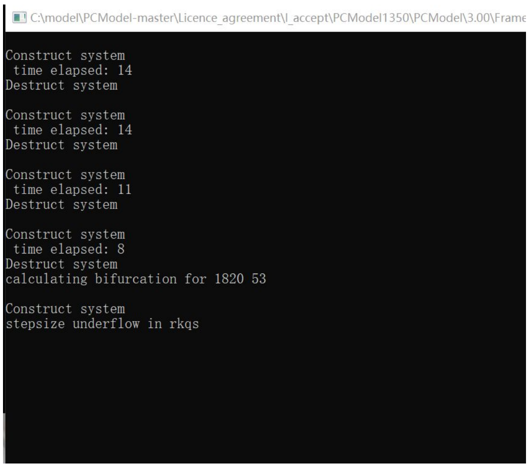
BIFURCATION

- Bifurcation run very short and destruct analyser



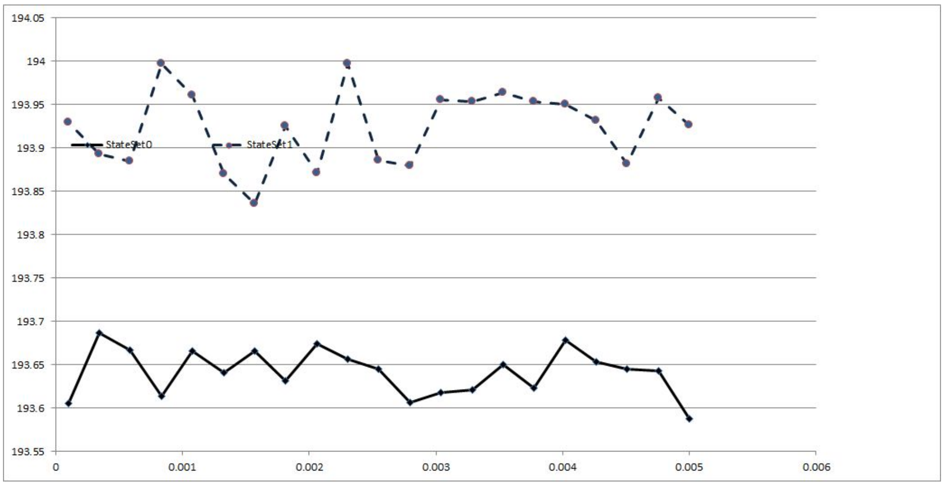
There was no parameter or derivative defined for the bifurcation analysis, Check column E of the 'parameter sheet' and the 'derivative sheet' and make sure that you put one 1 in the parameter sheet' and one 1 at the 'derivative sheet' in column E.

- Stepsize underflow in rkqs



If you get an "stepsize underflow in rkqs" you might have more than one 1's in column E of either the 'parameter sheet' or the 'derivative sheet'. Column E of both sheets can have only one 1.

- There is no bifurcation visible when running a bifurcation for mPload with oChla as output



If you get a graph like above, it is most likely that you did not put readPLoad at 1 in column J-M (depending on the one you used).

OPEN OUTPUT

- not directed to the right folder