

## **A command-line interface for automated simulations in SYMFOR**

**Paul D. Phillips**  
**2002**

Centre for the Study of Environmental Change and Sustainability (CECS)  
The University of Edinburgh  
John Muir Building  
Mayfield Road  
Edinburgh EH9 3JK  
United Kingdom  
Tel: +44 131 650 7860  
Fax: +44 131 650 7863  
Email: [info@symfor.org](mailto:info@symfor.org)  
<http://www.symfor.org/technical>

## Summary

---

SYMFOR is a framework for individual-based, spatially explicit models of the ecological and management processes affecting trees in tropical forests. It comprises a library of models and data storage facilities, accessed through an interface. Previously, the interfaces have been graphical, Windows-based interfaces designed mainly for people who use the framework occasionally, or are learning about it.

There are limitations associated with the graphical user-interfaces when applied to complex simulations such as may be performed by experienced users. These are related to the storage of run settings, the time taken to set up a multiple-run and the likelihood of making mistakes. These have been addressed by creating a new interface that operates through the “command prompt” window and a configuration file. Users modify a configuration file specific to the simulations they require, and the command-line interface interprets these and performs the simulations.

The command-line interface to SYMFOR is a DOS-type program written in C++ to execute under Windows operating systems. It is included with the latest (post September 2002) installations of SYMFOR, along with an example configuration file. Section 2 of this document serves as a manual for the program.

The command-line interface provides experienced SYMFOR users with a more convenient and efficient way to perform multiple-runs. It allows a more direct interaction between SYMFOR and statistical analysis packages that are optimised for different types of analysis. It also provides a pathway to adapt SYMFOR to execute under alternative operating systems, if this is required in the future.

The full text of this document is available from <http://www.symfor.org/technical>.

# Contents

---

<b>Summary</b> .....	<b>ii</b>
<b>Contents</b> .....	<b>iii</b>
<b>1 Introduction</b> .....	<b>1</b>
1.1 Introduction .....	1
1.2 SYMFOR program structure and code.....	1
1.3 SYMFOR user-interfaces .....	1
1.4 The need for a new interface .....	2
<b>2 The command-line interface: a manual</b> .....	<b>3</b>
2.1 Introduction .....	3
2.2 Simulations with the command-line interface.....	3
2.3 Structure of the configuration file .....	3
2.4 An example configuration file .....	5
2.5 An example tree data file list file .....	6
<b>3 Discussion</b> .....	<b>7</b>
3.1 The new UI.....	7
3.2 Usage of the command-line UI .....	7
<b>4 Conclusions</b> .....	<b>8</b>
<b>5 References</b> .....	<b>9</b>

# 1 Introduction

---

## 1.1 Introduction

SYMFOR is a framework for individual-based, spatially explicit models of the ecological and management processes affecting trees in tropical forests (Phillips and van Gardingen, 2001a; 2001b). It is implemented to run on personal computers utilising the Microsoft Windows™ operating system. In a typical simulation, data representing all the trees in an area are input to SYMFOR and the changes to the trees' state over a period of time is simulated.

Any one simulation uses a single Permanent Sample Plot (PSP) as input data, performs the simulation, and outputs either displays for visual examination, or files containing data produced in the simulation. A single PSP alone is not usually sufficient to represent the forest of which it is part. This is because of the dimensions of the PSP relative to the distance scale relevant to tropical forest processes and tree populations. In order to perform simulations that produce meaningful results over a larger area of forest it is necessary to use data from several plots in a series of simulations, considering each plot individually, and then combining the results.

Models within SYMFOR (Phillips *et al.*, 2002a; 2002b) often contain stochastic components, leading to variation in results when the same model is used with the same input data in repeated runs (with a different initial random number “seed”). It is necessary to perform repeated simulations on the same dataset in order to establish representative results. The command-line interface to the SYMFOR framework described in this technical note was implemented to assist this process.

## 1.2 SYMFOR program structure and code

The SYMFOR framework is implemented in two parts: a user-interface (UI), and a dynamic link library (DLL) that undertakes simulation processing and implements data storage. The code for these components is documented in detail by Phillips (2000).

The user-interface is a program that runs under the Microsoft Windows operating system, written in Visual Basic. It provides the user with suitable prompts of what data to input, what models to choose, etc. It presents the user with graphical displays, as requested, describing the data during a simulation. It also prevents the user doing many (but not all) things that would lead to an unsuccessful simulation. The UI communicates with the DLL to control the simulations.

The DLL is a software library written in C++, which awaits command from a Windows operating system executable. The DLL executes data input, data storage, model initialisation and execution and data output. It returns control to the UI program each simulated year, so that the UI can report to the user on the state of the simulation. The DLL also generates error messages as appropriate to be returned to the user. Any executable may call functions from the DLL, though it must be specifically written to do so.

## 1.3 SYMFOR user-interfaces

### Single-run mode

The single-run mode also goes by the names of “SYMFOR teaching” and “SYMFOR concise – single run”. It was designed primarily for demonstrations of the models, and presents the user with the minimum of options. The user can input data, select ecological and management models, choose to view displays, and run single simulations.

### Multiple-run mode

The multiple-run mode also goes by the name of “SYMFOR concise – multi-run”. It was designed for doing multiple simulations: repeated simulations on each of a set of plots. The user chooses the ecological and management models, the input data files, the number of repetitions, the length of the run and the output data details, and can then perform all simulations in one mouse-click. Only one combination of ecological and

management model can be used in any one multiple run. The same simulation settings are used for each input data file.

### Model Manager mode

The SYMFOR model manager is the original and most flexible of the UIs. With it, the user is able to select, view and modify the module choices and parameter values that make up the models. The user can also perform either single simulations or multiple simulations with the model manager, though with slight differences from the above two modes.

In a model manager single simulation, the user can change the models mid-run, can choose to output data, and can define new variables to be stored by the UI for temporal data displays.

In a model manager multiple simulation, the user must specify all the requirements of the simulation separately for each plot. It means that the simulation settings can be different for different plots within the same multiple-run. This is laborious and can be error prone. The user can store multiple-run “instructions” for later recall, editing or just for reference.

## 1.4 The need for a new interface

Usage of graphical user-interfaces is ideal for users who are not very familiar with the operation of a program. Experienced users attempting more complex simulations may find the Windows UI a distraction that is *more* likely to result in error.

This is the main reason for designing a new UI that requires minimal user interaction. There are some other reasons, however:

- The maximum number of plots usable in a simulation performed through the Windows multiple-run mode is 20.
- The maximum number of plots usable in a simulation performed through the model manager multiple-run mode is 200. This is not a practical limitation for most applications, but most users would find it preferable and more accurate to be able to automate the processing of large numbers of plots.
- The graphical UIs, written in Visual Basic, run more slowly than text-based or command-line driven software. This is not a *serious* problem, since the majority of processing is performed by the DLL, which is written in C++ and exhibits very fast execution.
- The Windows-based UI cannot be linked directly to other software packages such as statistical and GIS applications.

## 2 The command-line interface: a manual

---

### 2.1 Introduction

The SYMFOR command-line interface, “`symforc1.exe`”, is an executable program that interacts with the user through text in a DOS-type “command prompt” window. It is written in C++, and is compiled to execute under Windows operating systems. It is installed with the latest version of SYMFOR (<http://www.symfor.org>), and is found in the root SYMFOR directory.

The command prompt window is used only for the program to report progress and errors, and is never used for input. Input is performed through a configuration file, “`symforc1.cfg`”, located in the same directory as the executable.

This manual explains how to use the configuration file. It assumes some knowledge of the graphical SYMFOR UIs.

If you are still unsure how to use the graphical user-interfaces, you should not be trying to use the command line version!

### 2.2 Simulations with the command-line interface

The command-line interface to SYMFOR was designed for doing multiple simulations: repeated simulations on each of a set of plots. The user specifies in the configuration file the ecological and management models, the input data files, the number of repetitions, the length of the run and the output data details, and can then perform all simulations by running the program. Only one combination of ecological and management model can be used in any one command-line run. The same simulation settings are used for each input data file.

It is similar in function to the multiple-run mode described above (section 1.3), but is designed for simulations on many more plots. The maximum number of plots is currently set to 500, although this could be increased with re-compilation of the source code. The plots are listed in a separate file, which is pointed to by the configuration file. This has two advantages: the configuration file is kept simple; and the list of input files can be generated automatically by another program; for example, the program that creates the input files in the first place.

A copy of the configuration file and the file containing the list of input data files can be kept with other analysis and report files for a particular project, as an archive or logbook. This is much more convenient than the equivalent method for the graphical user-interfaces, where information about all the multiple-runs is stored in one file.

The easiest way to set up a configuration file is to edit one that already exists. This reduces the chance of making mistakes, and makes the process quicker.

### 2.3 Structure of the configuration file

The configuration file is a text file that should be edited with a text editor, such as “`notepad.exe`”, “`wordpad.exe`” or “`editpad.exe`”. Information is organised on lines, and the content of each line is important. The following sections refer to particular line numbers, or to blocks of lines in the configuration file.

#### Line 1

This is a comment line, for use by the user to identify the configuration file. It should begin with “//”, to signify that it is a comment, and should contain information that may be useful at a later date.

## Line 2 – 8

The structure of these lines is in the form “keyword=user choice”, where “keyword” refers to information required for a simulation, and “user choice” is the user’s specification of that information in a particular case. For example:

```
silvmodel=[RIL 40 m3ph 30 yr cycle]
ecolmodel=[Guyana-Pibiri]
nyears=160
nreps=10
inputfile=C:\Program Files\SYMFOR\runs\My Paper\My Paper filelist.txt
standfile=C:\Program Files\SYMFOR\data\stand\examplestand.csv
outputinfo=2                                2 output tables
```

These lines are probably self explanatory. They specify, respectively:

- the management (silvicultural) and ecological models in square brackets;
- the number of years for which each simulation should run;
- the number of repeat simulations to be carried out on each plot;
- the location of the file containing the list of input tree-data files;
- the location of the stand-data file (the same stand-data file is used for all tree-data files; this means the input tree data files must represent plots of the same size);
- the number of output tables that will be produced; the text “2 output tables” is a comment, and is ignored when the configuration file is interpreted.

Following the specification of the number of output files, the details of each output file in turn are required (if there are 1 or more output files).

## Output file details specification

For each file to be output, the following data are required:

- The file name and location to which to output.
- A flag (1 or 0) indicating whether to overwrite (0) or append (1) to a file, if one already exists with the same name and location (note that although a flag is required, it is ignored: the file is always appended to, so you should delete the old file first if you wish to ignore previous output data).
- Six flags (each 1 or 0), separated by spaces, indicating when to output data. a ‘1’ indicates that output is required, and a ‘0’ indicates that no output is required. These are in the following sequence:
  1. Output at regular intervals (for example every 10 years);
  2. Output before logging;
  3. Output after logging;
  4. Output at the start of the run;
  5. Output at the end of the run;
  6. Output at a fixed time following logging (for example 10 years after logging).
- Two numbers separated by a space; the first indicating the time interval used if “output at regular intervals” was selected; the second indicating the time interval used if “output at a fixed time following logging” was selected. These numbers may be followed by a text comment.
- The name of the SYMFOR forest object about which output data are required (for example, `livetree` or `gridsquare`).
- The number of columns required for the output table. This may be followed by a text comment.
- Lines defining the data to be output in each column. One line is used to describe data for each column. The lines contain the name of the column, followed by the name of the SYMFOR forest object attribute, followed by a flag that has the value ‘1’ for floating point numbers, and ‘2’ for integers. These entries are separated by “,” (comma space). Note that the names of variables are case-sensitive.
- This is followed by the end of the file, or the definition of another output table, depending on the number of output tables specified.

## 2.4 An example configuration file

The example configuration file, below, will run 10 simulations of 160 years on each of the tree data files contained in “...\\My Paper filelist.txt” with the “examplestand.csv” stand data, using the ecological model called “...\\Guyana-Pibiri” and the management model called “RIL 40 m3ph 30 yr cycle”. It will output two tables, one with data about `livetree` objects in SYMFOR, and one with data about `felledtree` objects.

`Livetree` data will be output to the file “...\\livetreecl.csv” every 5 years, at the start of the run and before and after logging, with the 8 columns: “PlotNo” (integer), “RepetitionNo” (integer), “DumpNumber” (integer), “speciesgrp” (integer), “utilgrp” (integer), “dbh” (float), “quality” (float) and “volume” (float).

`Felledtree` data will be output to the file “...\\felledtreecl.csv” at the end of the run, with the 5 columns: “PlotNo” (integer), “RepetitionNo” (integer), “yeardied” (integer), “speciesgrp” (integer) and “volume” (float).



```

// To examine heavy logging for the SYMFOR Brazil paper
silvmodel=[RIL 40 m3ph 30 yr cycle]
ecolmodel=[Guyana-Pibiri]
nyears=160
nreps=10
inputfile=C:\Program Files\SYMFOR\runs\My Paper\My Paper filelist.txt
standfile=C:\Program Files\SYMFOR\data\stand\examplestand.csv
outputinfo=2                                1 output tables
C:\Program Files\SYMFOR\result\livetreecl.csv
1
1 1 1 1 0 0
5 0                                regular intervals, interval following logging
livetree
8                                ncolums
PlotNo, PlotNo, 2
RepetitionNo, RepetitionNo, 2
DumpNumber, DumpNumber, 2
speciesgrp, speciesgrp, 2
utilgrp, utilgrp, 2
dbh, dbh, 1
quality, quality, 1
volume, volume, 1
C:\Program Files\SYMFOR\result\felledtreecl.csv
1
0 0 0 0 1 0
2 0                                regular intervals, interval following logging
felledtree
5                                ncolums
PlotNo, PlotNo, 2
RepetitionNo, RepetitionNo, 2
yardied, yardied, 2
speciesgrp, speciesgrp, 2
volume, volume, 1

```

This file could easily be edited to perform simulations on different data, using different models and producing different output. Information regarding the SYMFOR forest object attribute names can be obtained either from the SYMFOR online help files, or by using the SYMFOR model manager UI.

## 2.5 An example tree data file list file

The following is a valid tree data file list file. To add more files to the list, simply append them on the bottom, with one file per line. The program will continue reading the list until the end of the file is reached.

```

C:\Program Files\SYMFOR\data\tree\R1B1P1S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P2S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P3S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P4S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P5S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P6S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P7S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P8S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P9S1.csv
C:\Program Files\SYMFOR\data\tree\R1B1P10S1.csv

```

The files do not have to be in the same directory, but the full path and suffix (".csv") must be specified for each file.

## 3 Discussion

---

### 3.1 The new UI

The command-line interface to SYMFOR is suitable only for experienced users, however this expertise does not take long to acquire using the other UIs. Once a user is familiar with the concepts of SYMFOR and simulations using SYMFOR, the command-line interface provides a quicker and more reliable way to set up simulations, and a quicker and less resource-intensive way to run simulations. It allows the user to store simulation settings with the other files associated with a piece of work, rather than with the SYMFOR software files.

Potentially, the command-line interface paves the way for adaptation of the SYMFOR software to other operating systems, such as Linux or Unix. The command-line interface is a C++ program, written without use of any Microsoft-specific code. This means that it should compile under most C++ compilers, and therefore produce programs that can run on any operating system with a C++ compiler. On a different operating system, the command-line interface would still require a local SYMFOR DLL, however. The concept of a DLL, and many elements of the SYMFOR DLL source code, are specific to the windows operating system, and different versions would have to be produced to compile on different systems. This task is achievable, while the same cannot be said for the Visual Basic-based UIs traditionally associated with SYMFOR interfaces.

### 3.2 Usage of the command-line UI

The command-line UI was designed for SYMFOR simulations that use automatically generated tree data input files. This is a particular advantage when there are a very large number of input files that would be laborious to list manually.

One such case occurs in the process of preparing PSP data for calibration of a new model. Typically the data contain information that were recorded for each tree, but the growth model may use a competition index as a dependent variable. Thus the regression process would need a value of competition index for each tree in the PSP data, which is obtained by comparing each tree in a plot with each other tree in the same plot, and applying an equation. For a large dataset there may be 50 or more plots, that may have been enumerated 5 or more times, leading to over 250 plot-dates. The competition index is typically required for each of these enumerations. Fourth generation languages (4GL), such as are often used by databases or statistical analysis languages, are extremely inefficient at comparing each record with each other record in a dataset: the programs to do this are long, complicated and very slow to run. It is much quicker to execute the following process:

- output single-enumerations of each plot to a unique filename;
- make and output a list of the filenames;
- run SYMFOR on each file using the command-line interface UI, for 1 year outputting livetree data at the start of the run, including the competition index;
- importing the SYMFOR-output data into the analysis package;
- and integrating the SYMFOR-output data to the original tree data using the list of filenames.

This is a specific case where the command-line interface is essential. Many users will find that the interface is preferable to the graphical UIs for normal simulations as well, and will adopt it as their preferred method of conducting simulations.

The graphical UIs are not completely redundant, however. The model manager, in particular, is still needed for construction and alteration of models and for single runs to use the displays. The single run mode is still the UI of choice for demonstrations and presentations.

## 4 Conclusions

---

The command-line interface to SYMFOR is a DOS-type program written in C++ to execute under Windows operating systems. It is included with the latest (post September 2002) installations of SYMFOR, along with an example configuration file. Section 2 of this document serves as a manual for the program. It is the fourth user-interface to SYMFOR, which between them provide five interfaces in which to run simulations.

The command-line interface provides experienced SYMFOR users with a more convenient and efficient way to perform multiple-runs. It allows a more direct interaction between SYMFOR and statistical analysis packages that are optimised for different types of analysis. It also creates a pathway to develop implementations of SYMFOR to execute under alternative operating systems in the future.

## 5 References

---

- Phillips, P.D., van Gardingen, P.R., 2001a. The SYMFOR Framework for Modelling the Effects of Silviculture on the Growth and Yield of Tropical Forests. In the proceedings of IUFRO 4.11 Conference: 'Forest Biometry, Modelling and Information Science', (Editor; K.Rennolls), University of Greenwich, 25-29 June 2001. 12 pages. <http://cms1.gre.ac.uk/conferences/iufro/proceedings/>
- Phillips, P.D. and van Gardingen, P.R., 2001b. The SYMFOR framework for individual-based spatial ecological and silvicultural forest models. SYMFOR Technical Notes Series No. 8, The University of Edinburgh. <http://www.symfor.org/technical/framework.pdf>
- Phillips, P.D., van der Hout, P., Arets, E.J.M.M., Zagt, R.J., van Gardingen, P.R. 2002a. Modelling the natural forest processes using data from the Tropenbos plots at Pibiri, Guyana. SYMFOR technical note series no. 9, The University of Edinburgh, Edinburgh, 25pp. <http://www.symfor.org/technical/pibiri4.pdf>
- Phillips, P.D., Brash, T.E., Yasman I., Subagyo, S., van Gardingen, P.R. 2002b. An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo). Ecol. Model. (in press).
- Phillips, P.D., 2000. SYMFOR code documentation. SYMFOR technical note series no. 7, The University of Edinburgh, Edinburgh, 37pp. <http://www.symfor.org/technical/scd7.pdf>