

Outputting Simulations in Comma-Separated Values (CSV) File Format

The default file output formats used in DSSAT were designed to facilitate inspecting simulation results using a text editor or graphical tools specifically designed to work with the output files (e.g., GBuild and EasyGrapher). For applications involving large numbers of simulations, researchers usually analyze outputs using tools such as spreadsheets, scripting languages or statistical analysis software such as the R package (www.r-project.org) or SAS (www.sas.com). While routines can be written to read files using the default output formats with other software, a new option was introduced that allows users to request output files in CSV format (Fig. 1).

CSV format will be applied to EVALUATE.OUT and all *.OUT files that contain results as time series, while other files preserve existing formatting. To facilitate subsequent processing, each row additionally contains the run number, experiment identifier, treatment, rotation and replication numbers.

RUN	EXP	TRTNUM	ROTNUM	REPNO	YEAR	DOY	DAS	DAP	L#SD	GSTD	LAID	LWAD	SWAD	GWAD	RWAD	VWAD	CWAD	G#AD	GWGD	HIAD	PWAD	P#AD
1	PRPO0652	1	0	1	2006	314	41			0	0.000000						0	0.000000				
1	PRPO0652	1	0	1	2006	315	42			1	0.000000						0	0.000000				
1	PRPO0652	1	0	1	2006	316	43			2	0.000000						0	0.000000				
1	PRPO0652	1	0	1	2006	317	44			3	0.000000						0	0.000000				
1	PRPO0652	1	0	1	2006	318	45			4	0.1366245						0	0.5689302E-02				
1	PRPO0652	1	0	1	2006	319	46			5	0.4098562						0	0.6096334E-02				
1	PRPO0652	1	0	1	2006	320	47			6	0.6830686						0	0.7137947E-02				
1	PRPO0652	1	0	1	2006	321	48			7	0.9562619						0	0.8336549E-02				
1	PRPO0652	1	0	1	2006	322	49			8	1.243095						0	0.9737808E-02				
1	PRPO0652	1	0	1	2006	323	50			9	1.529910						0	0.1127783E-01				
1	PRPO0652	1	0	1	2006	324	51			10	1.816705						0	0.1275864E-01				
1	PRPO0652	1	0	1	2006	325	52			11	2.102492						0	0.1733314E-01				

Figure 1. Example of the PlantGro.csv file, showing only the first 12 columns of data.

The CSV output option can be specified using XBuild, the DSSAT experiment file creation tool, by modifying the “Simulation Options” for Outputs. Under the Format pull-down window, two choices are given (Fig. 2).

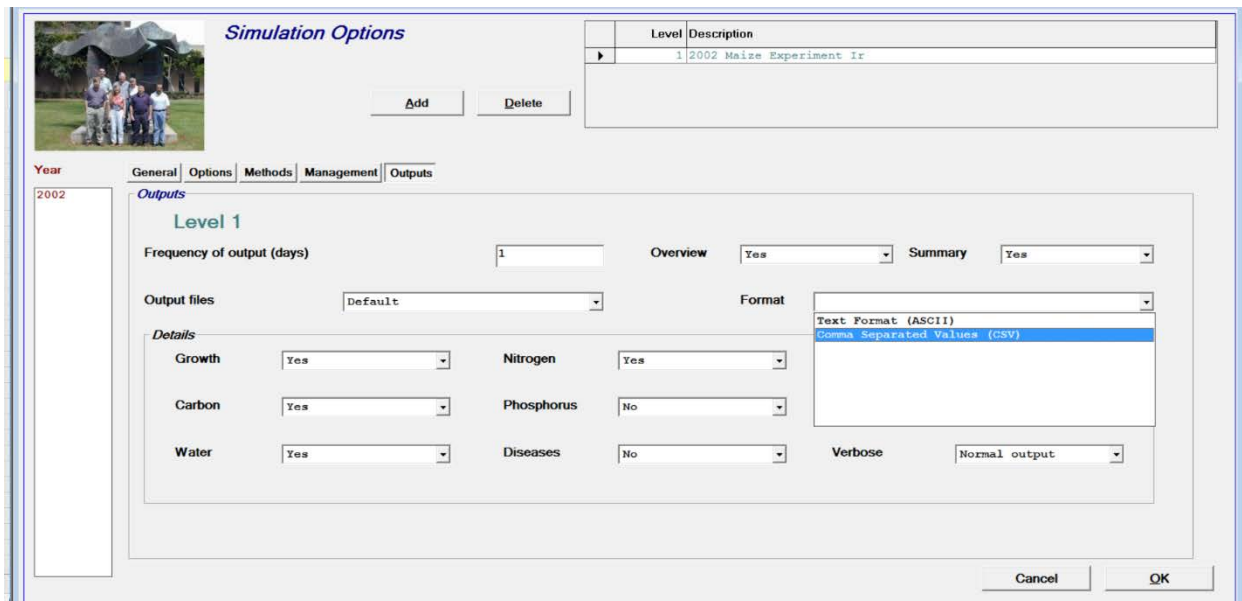


Figure 2. Example of the selecting CSV option using XBuild.

Alternatively the CSV option can be selected by setting the switch for the FMOPT variable to a value of “C” using a text editor. FMOPT is located under the OUTPUTS line of SIMULATION CONTROLS toward the bottom of a typical File-X (Fig. 3). (The value “A” results in traditional DSSAT .OUT formatting.)

```

11 06314 FE005 AP009      10  150      0      0      0      0      -99 -99

*SIMULATION CONTROLS
@N GENERAL      NYERS  NREPS  START  SDATE  RSEED  SNAME..... SMODEL
 5 GE           1      1      S 06274  2150  2006  WINTER
@N OPTIONS      WATER  NITRO  SYMBI  PHOSP  POTAS  DISES  CHEM  TILL  CO2
 5 OP           N      N      N      N      N      N      N      Y      M
@N METHODS      WTHER  INCON  LIGHT  EVAPO  INFIL  PHOTO  HYDRO  NSWIT  MESOM  MESEV  MESOL
 5 ME           M      M      E      R      S      C      R      1      G      S      2
@N MANAGEMENT  PLANT  IRRIG  FERTI  RESID  HARVS
 5 MA           R      R      R      R      M
@N OUTPUTS      FNAME  OVVEW  SUMRY  FROPT  GROUT  CAOUT  WAOUT  NIOUT  MIOUT  DIOUT  VBOSE  CHOUT  OPOUT  FMOPT
 5 OU           N      Y      Y      1      Y      Y      Y      Y      Y      Y      Y      Y      Y      C

@ AUTOMATIC MANAGEMENT
@N PLANTING     PFRST  PLAST  PH2OL  PH2OU  PH2OD  PSTMX  PSTMN
 5 PL           81090  81104  40     100    30     40     10
  
```

Figure 3. Portion of a File-X with the OUTPUTS line highlighted and the variable FMOPT to control output format indicated by a green arrow. The value "C" specifies that output be given in CSV format.

For the standard file naming option, specified by FNAME = N, files are named using the DSSAT convention for *.OUT files (e.g., PlantGro.CSV or Evaluate.CSV). Currently, if the filename option is selected (FNAME = Y), the option is ignored and files are still named using the *.OUT convention.

Because the CSV format does not require processing to convert each line of output to a specific field width and precision, using the CSV option can increase simulation speed by 10% or more. However, the CSV files will typically be 2.1 to 2.4 times larger than the corresponding *.OUT files.

An additional benefit of the CSV format is that it avoids problems with output values exceeding the allocated column width in the .OUT files. Figure 4 shows an excerpt from a maize simulation where an unrealistically high nitrogen input on DAS=31 resulted in overflow in NIAD (shown as “*****” in the SoilNI.OUT file). Furthermore on DAS=135, the variable NI1D was large enough to require the full eight columns allocated output, potentially causing problems for software expecting variables to be separated by a blank space.

```

!
!
! NO3 (ppm) by soil depth (cm):
! 0-5 5-15 15-30 30-45 45-60
@YEAR DOY DAS NAPC NI#M NLCC NIAD NITD NHTD NI1D NI2D NI3D NI4D NI5D
1999 365 0 0 0 0.0 5.8 5.8 0.0 0.50 0.50 0.43 0.40 0.27
...
2000 029 29 0 0 0.0 9.4 8.7 0.8 1.25 0.94 0.53 0.49 0.36
2000 030 30 0 0 0.0 9.6 8.8 0.8 1.27 0.95 0.54 0.49 0.37
2000 031 31 9999 1 0.0 ***** 5008.4 5000.2 3650.61 1825.60 0.54 0.49 0.37
2000 032 32 9999 1 0.0 ***** 6286.1 3712.7 4583.24 2299.12 0.54 0.50 0.38
...
2000 134 134 9999 1 0.0 9784.1 9783.5 0.6 9975.90 2145.42 0.46 0.82 0.74
2000 135 135 9999 1 0.0 9784.2 9783.6 0.610000.33 2133.20 0.46 0.82 0.74
2000 136 136 9999 1 0.0 9776.6 9776.0 0.610013.37 2121.06 0.47 0.83 0.75
  
```

Figure 4. Excerpt from a SoilNi.OUT file showing overflow (*****) and two variable fields that have run together.

The CSV option is fully implemented for CRGRO (22 crops), CSCER (Wheat, Barley), MZCER (Maize), SGCER(Sorghum), MLCER(Millet) crop models.

Other Restrictions:

1. The default output files often round values in order to fit data within the specified column width. The CSV format conserves the internal accuracy. This means that in comparing variables from the default files and the CSV files, values may differ slightly.
2. The CSV output is not compatible with the GBuild tool for graphical display. Our assumption is that people who use the CSV format will analyze their simulations with other (among them mentioned above) tools.
3. The CSV output option currently is not implemented for all crops. This is because output formats for different crops in DSSAT CSM were coded using different approaches. Crops currently not fully available include Sugarbeet, Barley, Cassava, Potato, Rice, Sugarcane, Taro, Tanier, Sweetcorn, Bermudagrass, Brachiaria, Bahiagrass, Perennial, and Alfalfa.
4. Files that are largely informational and have more complex formats are not available in CSV format. These include Overview.OUT and the various water and nutrient balance summaries.

Planned future developments:

Rather than expand the CSV format to other crops, ongoing work focuses on developing more flexible and generic input and output routines. This strategy will allow researchers to use data formats that best suit their needs and to limit output to those variables specifically of relevance for their simulation objectives.

Version notes:

The need for more “machine friendly” output formats was identified by various people using DSSAT in model intercomparisons coordinated by the Agricultural Model Intercomparison and Improvement Project (AgMIP). This led to collaboration between USDA ARS and the University of Florida from 2016 to 2017. The code changes were implemented by Vakhtang Shelia (UF) under the supervision of Gerrit Hoogenboom, with input from Cheryl Porter (UF) and Jeff White (USDA ARS), who also led testing. The first version was officially released in DSSAT 4.7.