Appendix: StateModB Input Type (StateMod Binary Output Files)

2004-07-27, Acrobat Distiller

Overview

The StateModB time series input type corresponds to the file format used by the State of Colorado's StateMod model, in particular the binary FORTRAN direct access output files. These files contain important water balance information for every node in the model network. The following table summarizes the contents of the binary files and corresponding text report files (all files can be large for large data sets):

Node Type	Monthly Binary File	Monthly Report File	Daily Binary File	Daily Report File
Diversion	**.b43*	**.xdd*	**.b49*	**.xdy*
Instream flow	**.b43*	**.xdd*	**.b49*	**.xdy*
Reservoir	**.b44*	**.xre*	**.b50*	**.xry*
Stream gage and Stream estimate	**.b43*	**.xdd*	**.b49*	**.xdy*
Well	**.b42*	**.xwe*	**.b65*	**.xwy*

The following documentation describes the format of the B43 binary file. Other files are similar. See the **StateMod Documentation** for a complete description of StateMod output files. Important comments about the file format are:

- The file is generated by StateMod as a direct access binary file with fixed-length records. The record length is 140 bytes.
- The file is divided into a header section (top) and data section (bottom).
- The format is optimized to allow a full year of data to be read for the entire data set. Efficiently reading a time series for a single location for the full period requires reading appropriate lines of the file using direct access. Because the file is binary and consistent for a given data set, file reads can be optimized.
- The data period and the calendar year type are consistent with the StateMod control file.
- All character strings are left justified and are padded with spaces. Therefore, software that reads the file should trim trailing spaces after reading the strings.
- River node identifiers in record 5 are included for all nodes in the network and data records (record 11) follow this order. Subsequent lists for various node types are a subset of the list in record 5 and have data items to reference the position in the river node list. Time series are queried using the identifiers in records 6+. However, the river node position is actually used to retrieve data in the file.

The B43 binary file contains the following records:

Record	Field	StateMod Variable	Type	Description
1	1	iystr0	integer	Beginning year of simulation, for year type in StateMod control file.
	2	iyend0	integer	Ending year of simulation, for year type in StateMod control file.
2	1	numsta	integer	Number of river nodes.
	2	numdiv	integer	Number of direct diversions.
	3	numifr	integer	Number of instream flows.
	4	numres	integer	Number of reservoirs.
	5	numown	integer	Number of reservoir owners.
	6	nrsact	integer	Number of active reservoirs.
	7	numrun	integer	Number of base flow nodes.
	8	numdivw	integer	Number of diversion structures with wells.
	9	numdxw	integer	Number of well only structures.
3	1	xmonam(14)	Each is char(4).	Month names corresponding to the calendar type for the simulation. This information is provided as a convenience for data processing. For example, if the year type is WYR (water year), xmonam(1) is 'OCT'. The 13th value is 'TOT' and the 14th value is 'AVE'.
4	1	mthday(12)	Each is integer.	Number of days per month, corresponding to the calendar type for the simulation. This information is provided as a convenience for data processing and to convert daily data values to monthly. For example, if the year type is WYR (water year), mthday(1) is 31 for October. The number of days in February is typically 28 and is used for all data processing, regardless of whether a year is a leap year.
5 **Repeat record for `numsta`**	1	j	integer	Counter for record type 5.
	2	cstaid(j)	char(12)	River node identifiers.
	3	stanam(j)	real(6)	River node names.
6 **Repeat record for `numdiv`**	1	j	integer	Counter for record type 6.
	2	cdivid(j)	char(12)	Diversion identifier.
	3	divnam(j)	real(6)	Diversion name.
	4	idvsta(j)	integer	River node position (1+) to allow cross-reference with river nodes.
7 **Repeat record for `numifr`**	1	j	integer	Counter for record type 7.
	2	cifrid(j)	char(12)	Instream flow identifier.
	3	xfrnam(j)	real(6)	Instream flow name.
	4	ifrsta(j)	integer	River node position (1+) to allow cross-reference with river nodes.
8 **Repeat record for `numres`**	1	j	integer	Counter for record type 8.
	2	cresid(j)	char(12)	Reservoir identifier.
	3	resnam(j)	real(6)	Reservoir name.
	4	irssta	integer	River node position (1+) to allow cross-reference with river nodes.

Record	Field	StateMod Variable	Type	Description	
9 **Repeat record for `numrun`**	1	j	integer	Counter for record type 9.	
	2	crunid(j)	char(12)	Base flow node identifier.	
	3	runnam(j)	real(6)	Base flow node name.	
	4	irusta(j)	integer	River node position (1+) to allow cross-reference with river nodes.	
10 **Repeat record for `numdivw`**	1	j	integer	Counter for record type 10.	
	2	cdividw(j)	char(12)	Well identifier.	
	3	divnamw(j)	real(6)	Well name.	
	4	idvstw(j)	integer	River node position (1+) to allow cross-reference with river nodes.	
11 **Repeat record for every river node `numsta`, for every month of simulation.** See the StateMod documentation for a full description of parameters. Parameters are grouped as shown in the **.xdd* file.	1	dat(1)	real	Demand	Total_Demand
	2	dat(2)	real	Demand	CU_Demand
	3	dat(3)	real	Water Supply	From_River_By_Priority
	4	dat(4)	real	Water Supply	From_River_By_Storage
	5	dat(5)	real	Water Supply	From_River_By_Exchange
	6	dat(60	real	Water Supply	From_Well
	7	dat(7)	real	Water Supply	From_Carrier_By_Priority
	8	dat(8)	real	Water Supply	From_Carrier_By_Storage
	9	dat(9)	real	Water Supply	Carried_Water
	10	dat(10)	real	Water Supply	From_Soil
	11	dat(11)	real	Water Supply	Total_Supply
	12	dat(12)	real	Shortage	Total_Short
	13	dat(13)	real	Shortage	CU_Short
	14	dat(14)	real	Water Use	Consumptive_Use
	15	dat(15)	real	Water Use	To_Soil
	16	dat(16)	real	Water Use	Total_Return
	17	dat(17)	real	Water Use	Loss
	18	dat(18)	real	Station In/Out	Upstream_Inflow
	19	dat(19)	real	Station In/Out	Reach_Gain
	20	dat(20)	real	Station In/Out	Return_Flow
	21	dat(21)	real	Station In/Out	Well_Depletion
	22	dat(22)	real	Station In/Out	To_From_GW_Storage
	23	dat(23)	real	Station Balance	River_Inflow
	24	dat(24)	real	Station Balance	River_Divert
	25	dat(25)	real	Station Balance	River_By_Well
	26	dat(26)	real	Station Balance	River_Outflow
	27	dat(27)	real	Available Flow	Available_Flow
	28	dat(28)	real	Structure type (Na): • < 0 = Baseflow node (e.g., -10001 indicates a diversion that is a baseflow node). • 0 = Well only. • 1-5000 = Diversion • 5001 – 7500 = Instream flow • 7501 – 10000 = Reservoir	
	29	dat(29)	real	Number of structures at this node (typically 1).	

StateMod B43 Files and Standard Time Series Properties

The standard time series identifier for StateMod binary time series is of the form:

```
Location.StateMod.DataType.Interval~StateModB~PathToFile
```

Time series properties are set using the following guidelines:

- The location part of the time series identifier is taken from the identifier field in the data. The identifier for the specific node type (e.g., diversion) is used, not the river node identifier. The river node identifier is often the same as for the specific node type, but this is not a requirement within StateMod.
- The data source part of the time series identifier is set to `StateMod`, because StateMod has created the output time series.
- The data type is assigned as the parameter name (see record 11 above, without using the group).
- The data interval is assigned as `Month` or `Day`, depending on the file extension.
- The scenario is set to blank.
- The input type is set to `StateModB`.
- The input name is set to the name of the file.
- The units for daily data are assigned as `CFS`. The units for monthly data in the files are average CFS for the month and are converted to `ACFT`, assuming a constant number of days per month, as read from record 4. February normally has 28 days per month in the header and therefore leap years have one fewer days than actual.
- The missing data value is assigned to `-999.0`.
- The description is set to the node name.
- The period is set based on the header information in record 1 (for the year) and record 3 (to determine the start and end months, based on the calendar type).

Limitations

StateMod binary files have the following limitations:

- The file does not contain a format version; therefore, it is difficult for software to handle changes in the file format.
- The file does not contain header information indicating the source of the file (e.g., the creation date, user, directory, StateMod response file, command line). Therefore, it is difficult to know with certainty how a file was created.
- Leap years are not explicitly handled with 29 days.
- Baseflow nodes in record 9 may have the same identifier as other nodes because any node can be a baseflow node. This can be confusing since software may list the node in more than one list. The software that reads the file filters out duplicate time series identifiers to try to resolve this problem.
- This documentation is limited in that it presents the file format only for the **.b43* file. Additional documentation may be added in the future.