

INPUT DATA STRUCTURE

This document describes a proposed format for instances of the stochastic hydrothermal unit commitment problem in electric power generation. Randomly generated, “realistic” instances encoded in this format are publicly available, and have been used to test some algorithmic approaches to the problem. In this section the structure is described and in the next on we give and illustrative example of the data in the files (all the data are text files with a .txt extension).

Five hydrothermal test systems are available for download, defined in Table I, considering different configuration of power plants, in terms of capacity and number of units. The hydroelectric power systems are based on real information of the Brazilian system. The hydroelectric data was extracted from a data base of the software HydroByte [1]. The data for the thermoelectric power plants was instead obtained from free (UC) instances, available at <http://www.di.unipi.it/optimize/Data> [2]. Finally, the transmission system and demand for each bus were adapted from an equivalent system from the south of the Brazilian electric power system. All the data is available online at <http://www.di.unipi.it/optimize/Data/UC.html>.

TABLE I
TEST SYSTEM FOR SHUC PROBLEMS

	Number of power plants		Generation Capacity (%)		Total Generation Capacity (MW)	Storage Capacity (hm ³)
	H	T	H	T		
A	7	14	25.0	75.0	21,297.5	5,635.1
B	7	14	75.0	25.0	9,224.0	9,309.0
C	10	10	50.0	50.0	16,132.2	10,737.5
D	14	7	76.3	23.7	16,046.5	14,944.1
E	14	7	25.2	74.8	9,671.0	5,507.2

H means hydroelectric and T means thermoelectric power plants. Storage capacity is the maximum volume to generate electricity.

Hydro plants (Hidreletricas.txt)

A row for each hydro plant is printed. The row is tab delimited and contains, in the following mandatory order:

1. A power plant index;
2. The minimum volume for the reservoir [hm³];
3. The maximum volume for the reservoir [hm³];
4. The minimum spillage for the reservoir [m³/s];
5. The maximum spillage for the reservoir [m³/s];
6. A flag that indicates the influence (1) or not (0) of the water spilled in the downstream level, in other words, indicates if the spillage changes the downstream level or not;
7. The number of groups of units;
8. An index of the upstream hydro plant;
9. The water time delay from the upstream hydro plant to the reservoir [h];
10. The columns regarding the numbers 10 to 14 are the coefficients of the upstream level function (a_0, a_1, a_2, a_3 and a_4), being the upstream level in metros equal to $a_0 + a_1v + a_2v^2 + a_3v^3 + a_4v^4$, where v is the volume of the reservoir in hm³;

15. The columns regarding the numbers 15 to 19 are the coefficients of the downstream level function (a_0, a_1, a_2, a_3 and a_4), being the downstream level in metros equal to $a_0 + a_1d + a_2d^2 + a_3d^3 + a_4d^4$, where d is the water discharged (given by the total water turbined and the spillage when the flag in the item 6 is 1 and given by just the total water turbined when the item 6 is 0) in the power plant in m^3/s .

Hydro initial conditions (Cond_Inic_H_X.txt)

In this case, the 'X' in file's name means variations of initial conditions, for instance, X can be 'a', 'b' and 'm', for high, low and medium (respectively) initial volumes.

A row for each hydro plant is printed. The row is tab delimited and contains, in the following mandatory order:

1. A power plant index;
2. The initial volume in the reservoir [hm^3];
3. The storage target for the reservoir [hm^3];
4. The future value of water stored in the reservoir [$R\$/hm^3$].

Hydro units (UnidadesH.txt)

A row for each group of units is printed. The row is tab delimited and contains, in the following mandatory order:

1. The power plant index;
2. A group index;
3. The number of units;
4. The minimum water turbined for each unit [m^3/s];
5. The maximum water turbined for each unit [m^3/s];
6. The minimum power output for each unit [MW];
7. The maximum power output for each unit [MW];
8. The coefficient of hydraulic losses k [$m/(m^3/s)^2$], the losses are given by kq^2 , where q is the water turbined in the hydro unit in m^3/s ;
9. The columns regarding the numbers 9 to 14 are coefficients of the turbine hydraulic efficiency function (a_0, a_1, a_2, a_3, a_4 and a_5), being the power output of each unit in MW equal to $a_0 + a_1q + a_2h + a_3qh + a_4q^2 + a_5h^2$, where h is the net head in m.

Thermal plants (Termeletricas.txt)

A row for each thermal plant is printed. The row is tab delimited and contains, in the following mandatory order:

1. A power plant index;
2. The minimum power output [MW];
3. The maximum power output [MW];
4. The minimum down-time [h];
5. The minimum up-time [h];
6. The maximum ramp down rate when unit is on [MW];
7. The maximum ramp up rate when unit is on [MW];
8. The columns regarding the numbers 8 to 10 are coefficients of the quadratic generation cost function (a_0, a_1 , and a_2), being the cost in $R\%$ equal to $a_0 + a_1pt + a_2pt^2$, where pt is the power output in MW;

11. The start-up cost can vary from a maximum "cold-start" value to a much smaller value if the unit was only turned off recently and is still relatively close to operating temperature. There are two approaches to treating a thermal unit during its down period, the start-up cost when cooling and when banking. The columns regarding the numbers 11 to 14 are the coefficients of the startup cost function (a_0, a_1, a_2, a_3), being the start up cost when cooling in R\$ equal to $a_0(1 - \exp(-x/a_2)) + a_1$, where x is the number of hours the unit was off before starting up. Also the start up cost when banking in R\$ is equal to $a_3x + a_1$.

Thermal initial conditions (Cond_Inic_T.txt)

A row for each thermal plant is printed. The row is tab delimited and contains, in the following mandatory order:

1. A power plant index;
2. The initial power output in $t = 0$ [MW];
3. The initial unit status, 1 indicates that the unit is on in $t = 0$ and 0 indicates that the unit is on off in $t = 0$;
4. The number of periods the unit was in the state defined in the item above.

Buses (Barras.txt)

A row for each bus is printed. The row is tab delimited and contains, in the following mandatory order:

1. A bus name;
2. The bus load index;
3. The hydro plant index;
4. The thermal plant index.

More than one power plant or demand can be considered for each bus, so when the bus index is the same in two or more lines indicates that more elements are included in this same bus.

Lines (Linhas.txt)

A row for each line is printed. The row is tab delimited and contains, in the following mandatory order:

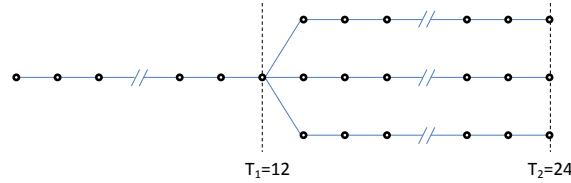
1. A line index;
2. The "from" bus index;
3. The "to" bus index;
4. The reactance of the line [pu];
5. The power flow capacity [MW].

General information (Parametros.txt)

Contains global information for the overall problem. The row is tab delimited and contains, in the following mandatory order:

1. The number of periods for the first stage, T_1 in the Figure below;
2. The total number of periods for the problem, or number of periods for the first and the second stage, T_1+T_2 in the Figure below;

3. The number of hours or fraction of hours for the first stage, in other words, the time discretization for the first stage [h];
4. The number of hours or fraction of hours for the second stage (from T_1 until T_2) [h];
5. The percentage of demand required for spinning reserve for each period [%];
6. The slack bus;
7. Non supplied power (deficit) cost [R\$/MW].



Load (Demanda_X.txt)

The 'X' in file's name means the number of scenarios in the file. For instance, X can be '4' and '9', for four and nine, scenarios of loads in the same file.

This information is divided for each scenario, so the data is divided in a set of rows, where a row for each load is printed. The row is tab delimited and contains, in the following mandatory order:

1. A load index;
2. The columns regarding the numbers 2,3,... are the load for each period considered [MW].

A set of lines is repeated for each scenario, for example, for a 2 buses system with 3 scenarios and 2 periods we could have:

1	1.00000E+02	9.00000E+01
2	5.00000E+01	6.00000E+01
1	1.00000E+02	1.10000E+02
2	5.00000E+01	4.00000E+01
1	1.00000E+02	1.10000E+02
2	5.00000E+01	6.00000E+01

The first and second lines are regarding the scenario 1, the third and fourth lines are regarding the scenario 2, and so on. The second column is regarding the time period $t = 1$ and the third column is regarding the period $t = 2$. Also this data is for a two-stage stochastic problem, so the value for the loads in the first stage ($t = 1, \dots, T_1$) need to be equal for all the scenarios.

Natural inflows (Afluencias_X.txt)

The 'X' in file's name means the number of scenarios in the file. For instance, X can be '4' and '9', for four and nine, scenarios of inflows in the same file.

This information is divided for each scenario, so the data is divided in a set of rows, where a row for each load is printed. The row is tab delimited and contains, in the following mandatory order:

1. A hydro power plant index;
2. A scenario index;
3. The scenario probability [%];
4. The columns regarding the numbers 4,5,... are the natural inflow for each period considered [m^3/s].

This set of lines is repeated for each scenario, for example, for 2 hydro plants with 3 scenarios and 3 periods ($T_1 = 1$ and $T_2 = 3$) we could have:

1	1	3.33333E+01	1.84000E+01	2.39000E+01	2.39000E+01
2	1	3.33333E+01	2.36000E+01	3.07000E+01	3.07000E+01
1	2	3.33333E+01	1.84000E+01	2.20000E+01	2.20000E+01
2	2	3.33333E+01	2.36000E+01	2.00000E+01	2.00000E+01
1	3	3.33333E+01	1.84000E+01	2.40000E+01	2.40000E+01
2	3	3.33333E+01	2.36000E+01	2.60000E+01	2.60000E+01

Also this data is for a two-stage stochastic problem, so the value for the loads in the first stage ($1, \dots, T_1$) need to be equal for all the scenarios.

EXAMPLE INSTANCE

We use a very small instance to illustrate the structure of the data. This small case is an electric power system with three hydroelectric power plants, two thermal ones, a transmission system with seven buses, ten lines and four loads. Further, the data consider a six hours scheduling horizon and two scenarios. The following Figures illustrate the power systems (on the left) and the scenario tree (on the right).

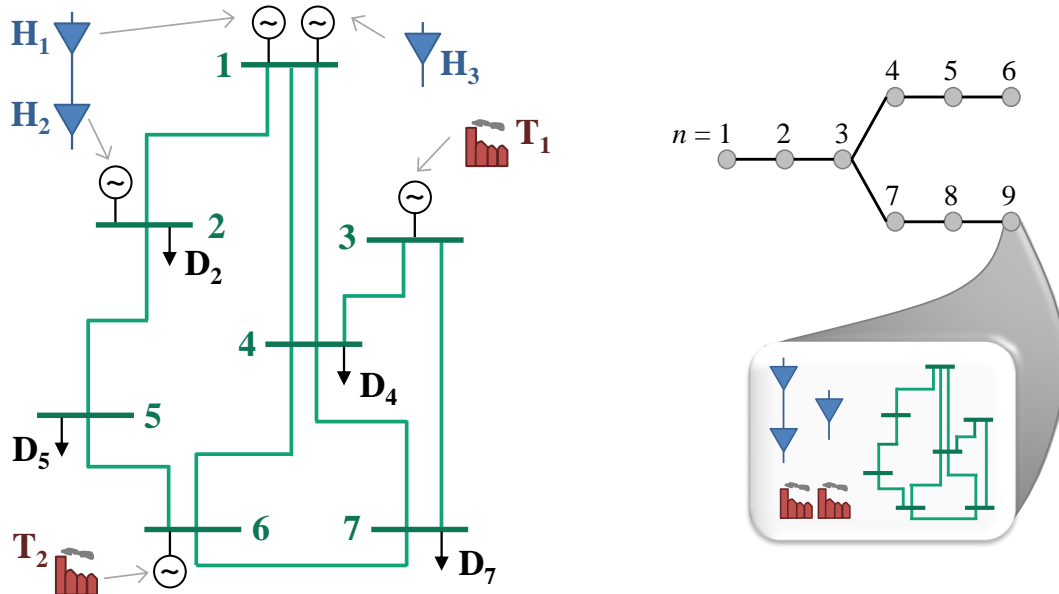


Figure 1. Power system and scenario tree for a very small instance

All the ten files are showed in the sequence as tables.

Hydro plants

Columns 1 to 9

H1	2.662000E+03	6.775000E+03	0.000000E+00	3.568000E+03	0	1	2	1
H2	9.480000E+02	1.124000E+03	0.000000E+00	4.040000E+03	1	2	0	0
H3	3.013000E+03	3.573000E+03	0.000000E+00	4.752000E+03	1	1	0	0

Columns 10 to 14 (coefficients of upstream level function)

4.477319E+02	1.823320E-02	-2.871219E-06	3.002598E-10	-1.272970E-14
3.660774E+02	4.485453E-02	-2.410526E-05	9.349116E-09	-1.447410E-12
-7.602610E+02	1.222976E+00	-5.234108E-04	1.001930E-07	-7.194833E-12

Columns 15 to 19 (coefficients of downstream level function)

3.944160E+02	2.111010E-03	-7.923070E-08	2.351570E-12	-2.713860E-17
3.218370E+02	2.280990E-03	-1.402710E-07	3.841880E-12	-5.363040E-17
2.579200E+02	6.208380E-04	-1.718400E-08	2.282680E-13	1.217070E-20

Hydro initial conditions

1	4.924150E+03	4.718500E+03	2.001240E+04
2	1.044800E+03	1.036000E+03	8.387550E+02
3	3.321000E+03	3.293000E+03	8.198617E+02

Hydro units

Columns 1 to 9

1	1	4	2.480000E+02	4.460000E+02	2.080000E+02	3.550000E+02	1.077585E-05
2	1	4	1.890000E+02	3.410000E+02	1.070000E+02	1.820000E+02	1.615610E-05
2	2	2	1.820000E+02	3.280000E+02	1.030000E+02	1.750000E+02	1.740499E-05
3	1	4	3.300000E+02	5.940000E+02	1.820000E+02	3.100000E+02	7.134364E-06

Columns 9 to 14 (coefficients of upstream level function)

3.587273E-01	2.638502E-03	3.444730E-03	8.700825E-06	-6.179360E-06	-2.83969E-05
3.587273E-01	3.451202E-03	5.136878E-03	1.697139E-05	-1.057229E-05	-6.31479E-05
3.587273E-01	3.589250E-03	5.136878E-03	1.765024E-05	-1.143499E-05	-6.31479E-05
3.587273E-01	1.937321E-03	5.372515E-03	9.963848E-06	-3.331440E-06	-6.90742E-05

Note: hydroelectric power plant H2 has two group of units, this requires two lines to describe its units, each line describes a group of units (minimum/maximum power output, hydraulic efficiency and so on).

Thermal plants

Columns 1 to 8

T1	5.700000E+01	1.060000E+02	2	3	1.540000E+02	1.410000E+02	6.087900E+02
T2	1.520000E+02	3.470000E+02	8	10	4.600000E+01	4.800000E+01	1.129110E+03

Columns 9 to 14

5.700000E+01	3.445560E-01	2.460000E+02	1.251000E+03	1.000000E+00	1.520000E+02
1.520000E+02	3.862300E-02	2.020000E+02	6.099000E+03	1.000000E+00	1.990000E+02

Thermal initial conditions

Columns 1 to 4

1	0.000000E+00	0	20
2	1.520000E+02	1	20

Buses

Columns 1 to 4

B1	1	1	0
B1	0	3	0
B2	2	2	0
B3	3	0	1
B4	4	0	0
B5	5	0	0
B6	6	0	2
B7	7	0	0

Note that bus 1 has two lines, the first adds the hydro plant 1 and the second adds (only) the hydro plant 3. Despite the bus 1 has no load, the first line receives an index for load. However, in the load data its value will be all zero.

Lines

Columns 1 to 5

L1	1	2	1.500000E-01	8.000000E+02
L2	1	4	2.300000E-01	1.000000E+02
L3	1	4	2.100000E-01	1.100000E+02
L4	2	5	1.600000E-01	6.000000E+02
L5	3	4	1.200000E-01	2.500000E+02
L6	3	7	1.500000E-01	2.500000E+02
L7	4	6	1.800000E-01	9.000000E+02
L8	4	7	1.900000E-01	9.000000E+02
L9	5	6	1.500000E-01	8.000000E+02
L10	6	7	1.500000E-01	6.000000E+02

General information

Columns 1 to 7

3	6	1.000000E+00	1.000000E+00	10	1	5.000000E+05
---	---	--------------	--------------	----	---	--------------

Load

Columns 1 to 7

1	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
2	5.000000E+02	4.950000E+02	4.990000E+02	5.030000E+02	5.070000E+02	5.100000E+02
3	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
4	4.500000E+02	4.530000E+02	4.500000E+02	4.420000E+02	4.400000E+02	4.450000E+02
5	4.000000E+02	3.940000E+02	3.900000E+02	3.900000E+02	3.870000E+02	3.870000E+02
6	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
7	3.000000E+02	3.100000E+02	3.150000E+02	3.150000E+02	3.100000E+02	3.050000E+02
1	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
2	5.000000E+02	4.950000E+02	4.990000E+02	5.050000E+02	5.090000E+02	5.130000E+02
3	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
4	4.500000E+02	4.530000E+02	4.500000E+02	4.410000E+02	4.390000E+02	4.420000E+02
5	4.000000E+02	3.940000E+02	3.900000E+02	3.870000E+02	3.870000E+02	3.820000E+02
6	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00	0.000000E+00
7	3.000000E+02	3.100000E+02	3.150000E+02	3.100000E+02	3.100000E+02	3.080000E+02

Columns 2, 3 and 4 represent the first stage periods, so they must be the same for both scenarios (scenario one, lines 1 to 7, and scenario two, lines 8 to 14). Columns 5 to 7 (three periods) represent the different realizations of the second stage, in this case the lines of scenario one and two could be different.

Natural inflows

Columns 1 to 6

1	1	5.00000E+01	8.920000E+02	8.820000E+02	8.780000E+02
2	1	5.00000E+01	1.010000E+03	1.015000E+03	1.020000E+03
3	1	5.00000E+01	1.188000E+03	1.190000E+03	1.190000E+03
1	2	5.00000E+01	8.920000E+02	8.820000E+02	8.780000E+02
2	2	5.00000E+01	1.010000E+03	1.015000E+03	1.020000E+03
3	2	5.00000E+01	1.188000E+03	1.190000E+03	1.190000E+03

Columns 7 to 9

8.800000E+02	8.760000E+02	8.780000E+02
1.010000E+03	1.013000E+03	1.009000E+03
1.188000E+03	1.188000E+03	1.188000E+03
8.800000E+02	8.820000E+02	8.830000E+02

1.012000E+03	1.009000E+03	1.000000E+03
1.200000E+03	1.200000E+03	1.200000E+03

Similar to the Load table, columns 4, 5 and 6 represent the first stage periods and columns 7, 8 and 9 the three periods of second stage. Columns 2 and 3 represent the index of the scenario and its probability, respectively, for each hydro plant.

REFERENCES

- [1] HydroByte, *HydroData XP*. <http://www.hydrobyte.com.br/site>, 2015.
- [2] A. Frangioni, *Generator of UC instances*. <http://www.di.unipi.it/optimize/Data/UC.html>, 2008.