1 Connections file `Connections.xlsx`

The file `Connections.xlsx` stores the directed water network in a structure designed by Alicia Forrester. The files stores the network in a 0-1 matrix with a structure similar to a (node-node) adjacency matrix (see, e.g., `http://en.wikipedia.org/wiki/Adjacency_matrix`). However, the structure differs from the standard adjacency matrix in the following ways:

1. The rows and columns are reversed from the adjacency matrix. I.e., if there is a 1 ith row and jth column, this indicates an arc from node j to node i.

2. The rows (referred to as *users*) and columns (referred to as *sources*) contain **only a subset** of the total nodes in the system.

   (a) Consequently, the ith row may represent to a different node than the ith column.

   (b) Thus, every node must be identified by a unique name.

      i. In the MATLAB code, this name must be a string.
      ii. Details on which nodes are included as a source is found in Section 1.2.
      iii. Details on which nodes are included as a user is found in Section 1.3.

   (c) Furthermore, this structure ensures that **some arcs are implied**, i.e., they exist in the network even though they are not shown in the matrix.

1.1 Node types

Every node is given a type, represented by a positive integer, describing its role in the water system. As of this writing, the possibilities are:

Type	Abbreviation	Number
Groundwater Source	GW	1
Waste Water Treatment Plant	WWTP	2
Surface Water Source	SW	3
Water Treatment Plant	WTP	4
Dummy Source	DMY	5
Recharge Facility	RCHRG	6
Potable Interceptor/Reservoir	P	7
Non-potable Interceptor/Booster	NP	8
Return Interceptor	RTRN	9
Potable Municipal User	PMU	10
Non-potable Municipal User	NPMU	11
Potable Industrial User	PIN	12
Non-potable Industrial User	NPIN	13
Potable Agricultural User	PAG	14
Non-potable Agricultural User	NPAG	15

1.2 Sources

The sources list contains every node with an outgoing arc, *except nodes where the outgoing arc is implied.* Every node that does not have a demand associated with it is listed as a source. Source

node names are given in row 1 of `Connections.xlsx`, beginning in column D. Row 2 contains the node type of each source.

Node types unique to Sources:

- Groundwater sources

- Surface water sources

- Dummy sources

1.3 Users

The users list contains every node with an incoming arc, *including nodes where the incoming arc is implied*. User node names are given in column A of `Connections.xlsx`, beginning in row 3. User (elevation) zones are given in column B. User types are given in column C.

Node types unique to Users:

- Potable users of all types (municipal, industrial and agricultural)

- Non-potable users of all types

1.4 Implied Outgoing Arcs of Potable Users

In a standard minimum cost flow problem, water enters the system at supply nodes and leaves the system at demand nodes. Non-potable demands in the water model behave in exactly this way. However, water satisfying potable demands is returned (after some loss) for treatment. Thus, we need to modify the standard formulation slightly to accommodate the fact that potable demand flows through the nodes.

Figure 1 shows a simple setup where two facilities supply potable water to a potable demand source, which then returns waste-water to a return interceptor. The problem needs a constraint ensuring that the total flow going *through* the demand node `DemP` meets the potable demand. Furthermore, we require that all water flowing into `DemP` be used to satisfy demand, i.e., the total flow into `DemP` exactly equals the demand. The code enforces these requirements by changing the system seen in Figure 2. The flow variables going into `DemP` are copied (subtracting the loss) and used as inputs in `RTRN`. Now, flow from `P` and `IPR` satisfy demand at `DemP`, given in the right hand side vector.
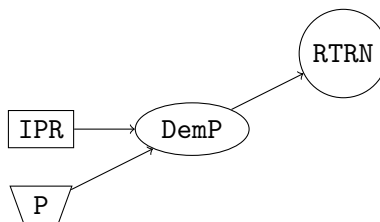

Figure 1: Schematic of two sources (`P` and `IPR`) supplying potable demand `DemP`, which return waste-water to `RTRN`. The arc connecting `DemP` to `RTRN` is implied, and is not specified in `Connections.xlsx`. For the implied arc to work, the nodes `DemP` and `RTRN` must be in the same zone.
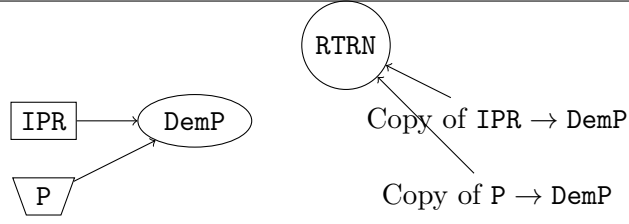
RTRN

IPR → DemP

P

Copy of IPR → DemP

Copy of P → DemP

Figure 2: Change of the network in Figure 1 to simulate demand flowing through potable demand node DemP.

Note: the arc to the return interceptor is implied, not stored explicitly in Connections.xlsx, and is required to be the unique return interceptor in the same zone as the potable demand.

2 Building the constraint matrix

2.1 Sub-matrices

The constraint matrix constructed by build_A.m covers all time periods of the problem, and is composed of three sub-matrices: A, A_st, and A_lag. Their basic functions are the following:

- A: Contains all variables in the current time period.

- A_st: Contains storage of currently usable water from previous time periods.

- A_lag: Controls the time required for water sent for treatment in previous time periods to become available for use.

Each of the three matrices can be further decomposed into sub-matrices. For example, A can be decomposed as

$$
\mathtt{A} = \left(\begin{array}{cccc} A & S & R & -I \\ A_{\mathrm{ws}} & S_{\mathrm{ws}} & R_{\mathrm{ws}} & 0 \\ L & 0 & 0 & I \end{array} \right),
\tag{1}
$$

where I is the identity matrix of appropriate size and 0 are the matrices of zeros of appropriate size. The submatrices have dimensions given in Table 1:

rows \ columns	numArcs	numStorage	numRelease	numUsers
numUsers	A	S	R	$-I$
numWS	A_{ws}	S_{ws}	R_{ws}	0
numUsers	L	0	0	I

Table 1: Size of each of the sub-matrices in (1).

The dimensions are:

- numUsers: Number of nodes listed as users.

- numArcs: Number of arcs defined in Connections.xlsx (i.e., not including implied arcs).

- `numStorage`: Number of nodes (groundwater sources and recharge facilities) capable of storing water over time periods.

- `numRelease`: Number of nodes (surface water sources and waste-water treatment plants) that can release water into the environment.

The purpose of each sub-matrix can be determined by it's row and column. The purposes of the rows:

- The first row stores all variables for each node represented as a user in `Connections.xlsx`.

- The second row stores all variables for each water source node (groundwater or surface water), which are not given as users in `Connections.xlsx`.

- The third row stores the water lost as it comes into each node listed as a user.

The purpose of the columns:

- The first column is for flows along arcs.

- The second column is for storage decisions.

- The third column is for releasing water back into the environment.

- The final column connects the movement of water to its transport losses.

In sub-matrix `A_st`, only the storage columns are nonzero. Likewise, in sub-matrix `A_lag`, only the arc columns are nonzero. In both `A_st` and `A_lag`, the identity matrices are replaced with zeros.

2.2   Combining the sub-matrices for year-long time periods

These sub-matrices can be combined in various ways depending on the desired time-step for each period. The variable `timeLag` controls the number of time steps per year. When `timeLag = 1`, the sub-matrices fit together in a block bi-diagonal form, which for 4 time periods looks like

$$\begin{pmatrix} \texttt{A} & & & \\ \texttt{A\_st} + \texttt{A\_lag} & \texttt{A} & & \\ & \texttt{A\_st} + \texttt{A\_lag} & \texttt{A} & \\ & & \texttt{A\_st} + \texttt{A\_lag} & \texttt{A} \end{pmatrix}.$$

2.3   Constraint Matrices for other time lags

As the code is currently written, the only options are `timeLag = 1` (for 1-year time steps) and `timeLag = 12` (for 1-month time steps). For illustration purposes, the constraint matrix is constructed as follows for `timeLag = 3`

$$\begin{pmatrix} \texttt{A} & & & & & \\ \texttt{A\_st} & \texttt{A} & & & & \\ & \texttt{A\_st} & \texttt{A} & & & \\ \texttt{A\_lag} & & \texttt{A\_st} & \texttt{A} & & \\ & \texttt{A\_lag} & & \texttt{A\_st} & \texttt{A} \end{pmatrix}.$$

That is, the sub-matrix `A_lag` is `timeLag` columns to the left of sub-matrix `A`. Note that for `timeLag = 1` the matrices `A_st` and `A_lag` will overlap, and thus are added together.

2.4 Loss Factors and Variables

Loss variables, in the L sub-matrix of `A` and `A_lag`, store the amount of water lost as the flow travels from the source to the user. Loss variables (and constraint rows) are named after the user nodes, with a suffix `-Loss` or `-loss` added, depending on whether the loss is the standard amount or not. The standard amount of loss is 1%, and is applied to all arcs by default. Exceptions are as follows:

1. Arcs leaving dummy nodes are not subject to loss.

2. Arcs coming into a recharge facility from either a waste-water treatment plant or a surface water source are not subject to loss.

   (a) Note: these are the only types of arcs entering a recharge facility in my connections matrices. Perhaps the code could be changed to have no loss for any arcs entering a recharge facility.

3. Arcs coming into a waste-water treatment plant from either another waste-water treatment plant or a water treatment plant have a 5% loss to solids.

   (a) Note: Return interceptors can also connect to waste-water treatment plants, and these arcs are subject to the standard loss.

4. Arcs coming into recharge facilities from previous time periods have a 3% loss due to evaporation.

There is a potential instability in the code here. Loss variables are characterized as standard or not based on the user node. However, some users (WWTPs and potable/non-potable users) may have arcs of both standard and non-standard loss coming into them. This may make the classification of some loss variables arbitrary. This may not be a problem, but it should be kept in mind.

3 Inputs File `Inputs.xlsx`

The second file, containing the cost vector, variable lower and upper bounds, and right hand side vector, is `Inputs.xlsx`. The file is largely self-explanatory, but there is one important catch: **There is no consistency checking in the MATLAB code**. The variables and constraints must be specified in exactly the order that they are read in MATLAB, because the code does not compare the internal names of the variables with the names supplied in the spreadsheet.