APPENDIX IV

WATER QUALITY MODELING DETAILED PROCEDURES AND TOOLS

WATER QUALITY DATABASE

A database is created for each *River GeoDSS* Base-Network to store the processed water quality data and the Water Quality Module (WQM) user modeling preferences. The database is implemented in MS-Access as an efficient alternative to store object oriented MODSIM-native time series (time series object type used in the WQM). The database uses the MODSIM Network name suffixed by "*TS*" with the MS-Access files extension (*mdb).* The database stores the network object user entered data in individual database tables. These data tables are named with the object MODSIM name. Two type of data tables are found in the database: (1) Concentration data tables and (2) flow-concentration fitted equation coefficients. Concentration data tables are suffixed with *_CONCENTRATION* and contain two columns: (1) the simulated period start date and (2) the concentration value. The fitting equation coefficients tables are suffixed with *_FITTINGCOEFS*. This single-column table contains in its rows the equation coefficients. A modeling preferences support table is stored in the database with the name *NodesWQCalibrationInfo*. This table contains water quality modeling for all the nodes in the network. *NodesWQCalibrationInfo* contains: (1) a flag (*USEQCFunction*) indicating if an equation is used to predict concentration as a function of flow, (2) the type of flow-concentration relationship equation

in column *CurveType*, (3) the minimum and maximum (*Min_Bound* and *Max_Bound*) values for calibration flows concentration and (4) the data filter used for the fitting equation calculation dataset. Figure IV-1 shows an example of the internal database table structure.

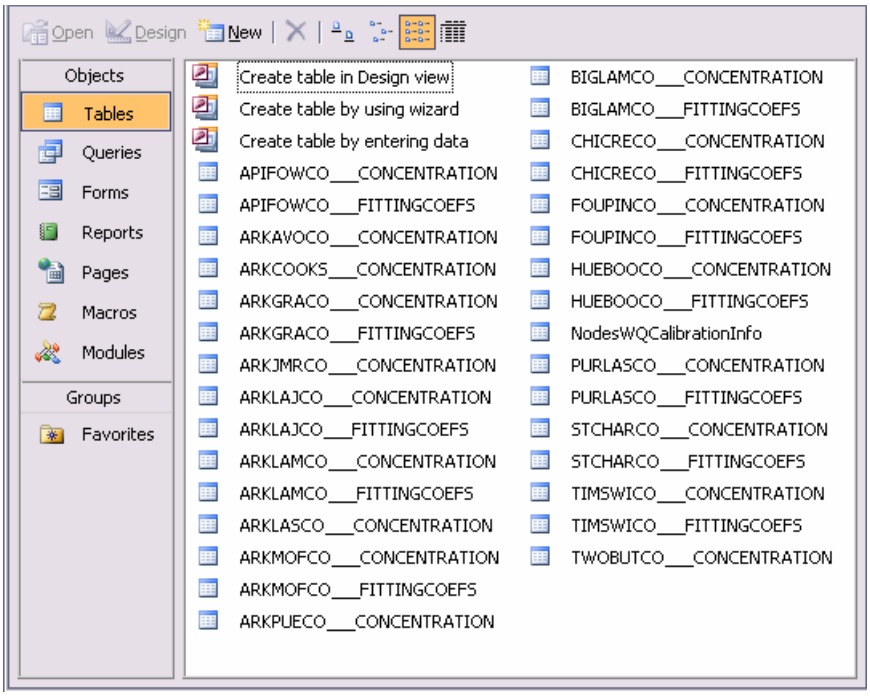

Figure IV-1 – WQM database tables sample

NETWORK TRACING ALGORITHM

The algorithm's basic loop goes through all the nodes in the network finding the terminal nodes, identified as a demand or sink node with no real links flowing out the node. For each of the terminal nodes, a nodes collection is built navigating upstream in the network through all the connected nodes towards the source nodes (identified by not having links coming into the node). The node collections supplies all nodes visited from a sink to a source node. These collections are combined into the final sequence using the following rules:

If the node to be added to the collection already exists, the node is not added.

If the node to be added is found in one of the final sequences, the member of the currently processed collection are inserted into the final sequence at the location where the node was found. The collection is then cleared and the node is not added to the collection.

The following diagrams (Figure IV-2 and Figure IV-3) illustrate the procedure implemented to trace the MODSIM network. *TODO* is a collection object that stores nodes to be processed, ← implies object code assignment, and *Final Order* is the final traced node sequence.

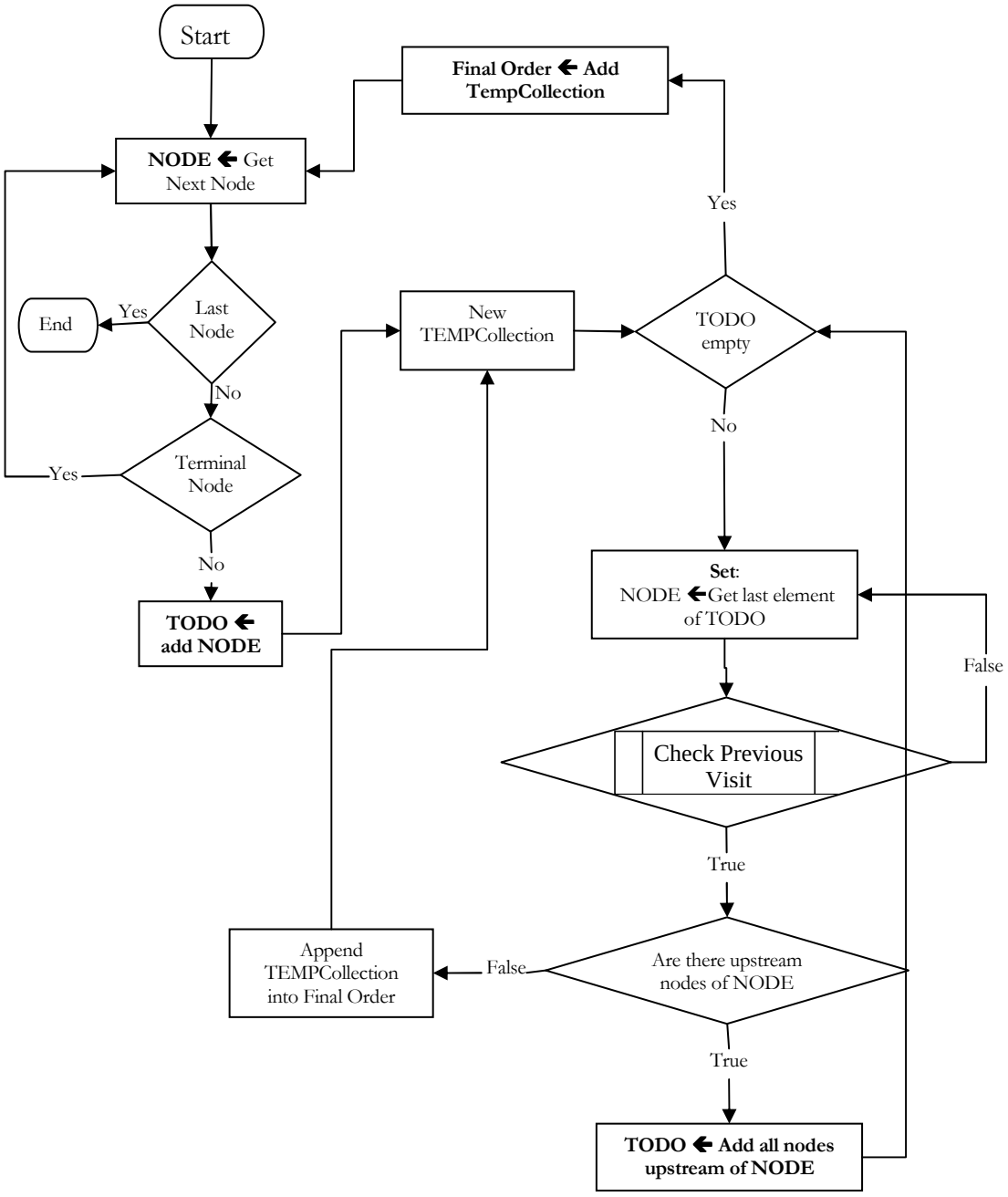

Figure IV-2 – Flowchart of the network tracing algorithm

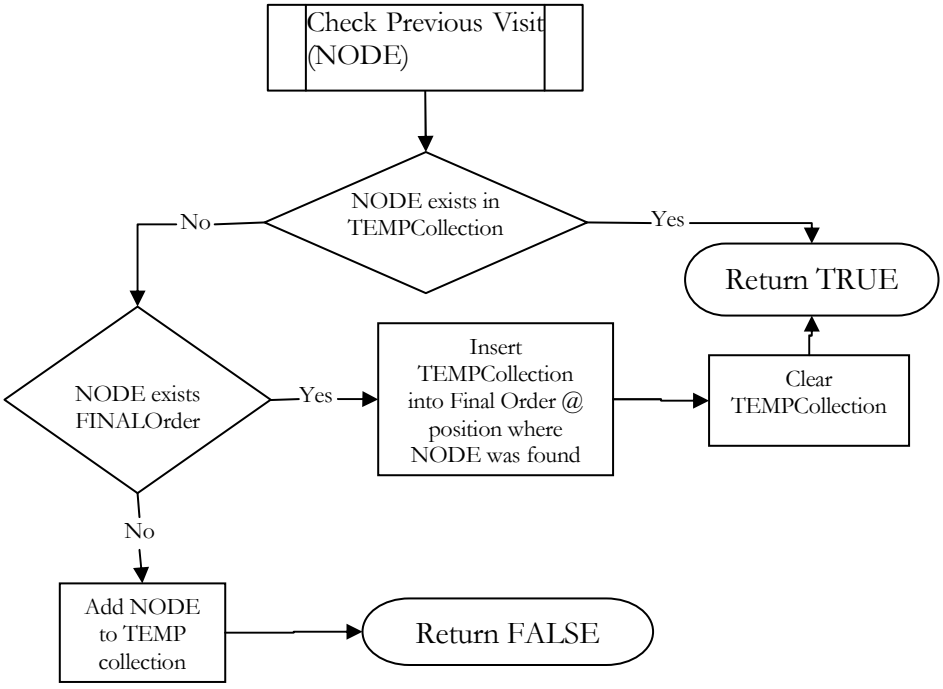

Figure IV-3 – Flowchart of the function Check Previous Visit

CONSERVATIVE CONSTITUENT ROUTING CALCULATION

The algorithm developed use the principle of conservation of mass at the system nodes to calculate concentrations in all the network links. At this stage, the algorithm is applied to conservative constituents such as the total dissolved solids (salt routing).

Inflow Concentrations and Measured Control Points

The MODSIM node Tag object is used to implement a water quality class that contains: (1) a variable that implements the time series MODSIM Class and (2) a variable size array that resembles the MODSIM model time series variables populated when the model is initialized with the processed values for all the model time steps.

In the modeled basin, Total Dissolve Solids (TDS) is inferred from the measured Specific Conductance and imported to the corresponding network nodes. Water concentration is required for all system water sources for the WQM mass routing algorithm. The unknown concentrations are assigned during the calibration procedure to minimize the deviation between the calculated and measured concentration. The WQM allows the user assigning concentrations to inflows at any place in the system. The algorithm assigns *NonStorage* node-entered concentrations to the node's inflow link. In addition, demand nodes acting as source nodes (connected to a water source structure) with measured concentration will be implemented assigning concentration to the flow in the WQM support links suffixed with "_CALIB_SOURCE". Undefined values will be assigned with zero concentration and warning message will be generated (This should be avoided because it will change significantly the modeled concentration). When the measured concentration is not located at a system inflow, the measured water concentration is used as calibration check point to compare the calculated concentration with the measured one. The algorithm is designed to identify the calibration control points based on the node concentration definition and the non-existence of inflow links.

The basic routing algorithm assigns the measured concentration to local gains at gauging stations with measured concentrations. In these cases, the measured concentration is assigned to the WQM support link named with the station name and the suffix "CALIB_DS_SUPPLY".

Tributaries and River Convergence Modeling

Aquifer return flow concentration from tributaries and the main stream could be different for the links entering a convergence node. At this point, there are multiple links upstream

and only one link supplying return flow; therefore, the single link concentration needs to be computed as a function of the individual concentrations. The salt load calculation at those junctions is carried out after MODSIM has converged to a solution, in which the link conveying the salt load to the convergence point is assigned with the combined return flow (tributary + Main stream). The ANN predicted salt mass load is operated with the total flow in the return link to set the concentration in the link. The resulting concentration (C) of combining tributary and main stream return flow and mass load is:

$$C = \frac{m_m + m_t}{V_T} = \frac{m_m}{V_T} + \frac{m_t}{V_T} = C'_m + C'_t \qquad \text{(IV-1)}$$

where, m_m = the mass load in the main river, m_t = the mass load in the tributary and V_T = the total combined water volume returned to the surface system. C'_m and C'_t are intermediate concentration calculations using the corresponding masses but the combined volume. During the water constituent routing calculation process, one of the intermediate concentration (C') will be initially calculated, the second intermediate concentration will be added to compute the final concentration in the link.

Water Concentration Calculation

The calculation of the constituent concentration for each link in the system is performed after the MODSIM solver has converged to a solution. The solution provides the water flow in all the links in the system. The concentration is sequentially calculated from upstream to downstream. The node calculation order assures that all the preceding concentrations are known when calculating the output concentration at any point in the system. Full mix of the constituent is assumed at each node. In consequence all the outflow nodes will have the same concentration.

The Tag object of the MODSIM links is used to carry a class that holds the concentration value for the current time step. The artificial inflow link is the only artificial one that is designed to have concentration assigned. The concentration in this link is populated each iteration based on the available user defined node-concentrations.

ANN interaction

Quantity and quality predictions of the return flow from the aquifer to the stream can be obtained in the *RIVER GEODSS* using the ANN module. The module uses a previously trained ANN to estimate the return flow water salinity. The concentration of the water added to the system from the groundwater is assigned to the links that provides linkage between the ANN source and the nodes in the system. These links are "real links", their concentration is assigned during the MODSIM solution iterative to the concentration variable in the link Tag object.

The ANN predicts total salt load to the stream, therefore the concentration is calculated as the salt mass divided by the water volume. The concentration units are converted to kg/m^3 or 1000kg/AcF in Metric and English systems respectively.

IMPLEMENTATION OF THE WATER QUALITY CALIBRATION

When the network is initialized, the algorithm identifies the network reaches upstream of all the station that have water quality data. The reaches (control volumes) are defined between water quality measuring stations and are composed of all the nodes that potentially contribute with mass to the downstream station's concentration. The algorithm trace the network upstream marking the nodes found until it reaches a demand node (that is not gauging station) or a demand node that is a gauging station and measures water quality.

During the network tracing, nodes with measured inflow, known concentration and unknown concentration are identified. The algorithm uses a table in memory to keep the relationships between each node and its corresponding downstream gauging station, as well as, the contributions from each node to the station's concentration calculation. A row is created for each node, the field *DWSStation* is used to link a node with the corresponding water quality station downstream. The table field *TypeID* is used to categorize the nodes. Initially all nodes are flagged with *TypeID* = 0. Other fields in the table are used to store mass added to and subtracted from the reach at each node. The *TypeID*=1 indicates that the node is brings a load to the reach with no concentration defined. The *TypeID*=2 indicates that is a calibration inflow to the reach. Calibration flows represent a combination of water from different sources within the reach, each one with different unknown concentrations. Initially, this calibration flow concentration is approximated with the concentration at the station upstream of the reach where calibration water is being provided but it will be adjusted later in the calibration if need it. *TypeID*=3 is used to identify nodes with measured concentration (gauging stations) or unchangeable-calculated concentration (e.g., terminal interfaces, ANN aquifer contributions). The table entries include columns to hold the flow in/out the nodes and their calculated output concentration. For each node, the algorithm calculates the total mass in and out of the reach. The mass balance terms at the nodes include: (1) reach external sources, (2) reach internal mass transports and (3) reach external sinks. Using positive (+) sign for all inputs and negative (-) sign for all the outputs, the internal mass transport will cancel out when combining all the nodes in the reach, leaving only external sinks and sources.

An iterative procedure is implemented to attempt to calibrate the model to measured concentrations. At the gauging stations, the difference between the simulated and the measured concentration is calculated. In cases where there is a discrepancy, a concentration (the same at all the unknown points) is calculated. The calculation is based on (1) the known concentration and inflows, (2) the calibration inflows using the upstream node measured/calculated concentration and (3) the inflow at the unknown concentration nodes. Equation 3.3 evaluated for the reach control volume may be written as:

$$\sum_{i=1}^{N} Q_i \cdot C_i - \sum_{j=1}^{M} Q_j \cdot C_j = 0 \qquad \text{(IV-2)}$$

where, N = the set of links with conveying mass into the reach, Q_k = the flow rate at link k of the reach, and C_k = concentration of water in link k. M = the set of links moving mass out the reach. Separating known and unknown concentration terms the mass conservation equation is:

$$Q_{out} \cdot C_{out} = \sum_{KnownC} Q_i \cdot C_i + \sum_{UnknownC} Q_i \cdot C_i - \sum_{j=1}^{M'} Q_j \cdot C_j \qquad \text{(IV-3)}$$

where, Q_{out} = the flow at the downstream end of the reach and C_{out} = the measured concentration at the downstream end of the reach. M' = the set of outflow links excluding the reach downstream station outflow link. Rearranging terms,

$$Q_{out} \cdot C_{out} - \sum_{KnownC} Q_i \cdot C_i + \sum_{j=1}^{M'} Q_j \cdot C_j = \sum_{UnknownC} Q_i \cdot C_i \qquad \text{(IV-4)}$$

Assuming a constant concentration C for all the unknown concentration inflow nodes,

$$\overline{C} = \frac{Q_{out} \cdot C_{out} - \sum_{KnownC} Q_i \cdot C_i + \sum_{j=1}^{M'} Q_j \cdot C_j}{\sum_{UnknownC} Q_i} \quad \text{(IV-5)}$$

The value $\overline{C}$ is to be constraint between observed concentrations in the area to be valid. The range of values for each node is selected in the water quality user dialog (Figure 3.12). Initially for each reach, the nodes with *TypeID* = 1 are assigned with a concentration to match the measured concentration downstream (Equation IV-5). In the first concentration modeling, the unknown concentrations were assumed as 0, therefore the average concentration for all the unknown sources is calculated as:

$$\overline{C} = \frac{\sum_{TypeID=0(OUT)} Q_j \cdot C_j - \left(\sum_{TypeID=0(IN)} Q_i \cdot C_i - \sum_{TypeID=1} Q_i \cdot C_i \right)}{\sum_{TypeID=1} Q_i} \quad \text{(IV-6)}$$

Notice that the measured flow and concentration at the downstream station are included in the sum of the TypeID =0 (OUT). When no nodes with TypeID = 1 are found, the concentrations in the nodes with TypeID = 2 are adjusted. These nodes were considered having known concentrations in the initial sum of mass in the reach therefore they need to be removed from the *known* sum.

$$\overline{C} = \frac{\sum_{TypeID=0(OUT)} Q_j \cdot C_j - \left(\sum_{TypeID=0(IN)} Q_i \cdot C_i - \sum_{TypeID=2} Q_i \cdot C_i \right)}{\sum_{TypeID=2} Q_i} \quad \text{(IV-7)}$$

The outflow concentrations are a function of the current inflow concentrations. Once the inflow concentration has been adjusted, the outflow concentration change; therefore, an iterative procedure is implemented to recalculate outflow concentrations and adjust

unknown concentrations accordingly until the calculated inflow concentration converges to the same concentration in consecutive iterations.

When groundwater quality modeling is active in the calibration, the ANN predicted values are net returns to the river. In these cases, the salt load to the river is treated as an input to the reach. There are few cases in which the net return is negative, indicating a removal of mass from the reach. The algorithm is implemented so that the net depletion from the river is accounted in the mass leaving the control volume (Calculated concentration at depletion point is assumed).

Convergence Criteria

The algorithm performs a set of checks to stop the iterative calculation of unknown concentrations at all the water quality stations. At any water quality station, the unknown concentrations are recomputed if (1) the difference between the calculated and measured concentration is larger than one unit, or (2) if the previous iteration calculated concentration and the current concentration doesn't change in more than 0.01 mg/L. In addition, the iteration process is limited to 60 iterations. After the 60^{th} iteration the calibration process attempt is dropped and the latest concentrations are used to calculate concentration downstream of the control point. For each time step, the calibration process finishes with an additional iteration in which all the calculated concentrations are use to route the constituent throughout the system.

Tributary Return Flow Special Consideration

Tributaries return flow and salt load are predicted in the mayor tributaries, where continuous water flow was observed in the field. Calibration on tributaries with gauging

stations need to account for the upstream contributions to set the concentration of the source link that provides the calibrated flows to match historical measured flows. The following adjustment takes place while setting the measured concentration to the station's source link. The conservation of mass at the node can be written as:

$$\sum_{IN} V_i C_i = \sum_{OUT} V_j C_j \tag{IV-8}$$

where, *IN* = the set of links entering the node and *OUT* = the set of links leaving the node. V_i = the volume of the link *i* from *IN*. C_i = the constituent concentration for link *i*. V_j = the volume of the link *j* from *OUT*. C_j = the constituent concentration in link *j*. Separating the source link terms (V_s and C_s) from *IN* and the measured terms (V_m and C_m) from *OUT* the mass balance equation at the node can be expressed as:

$$V_S C_S + \sum_{IN'} V_i C_i = V_m C_m + \sum_{OUT'} V_j C_j \tag{IV-9}$$

where, *IN'* is the set of incoming links excluding the source link. In a similar fashion, *OUT'* is the set of outgoing links excluding the stream link downstream of the station (with measured flow and concentration). The concentration in the source link will be calculated as:

$$C_S = \frac{V_m C_m + \sum_{OUT'} V_j C_j - \sum_{IN'} V_i C_i}{V_S} \tag{IV-10}$$

Additional handling is required for joints where tributary links and main river links blend together. Each of the links will have a predicted return flow and concentration. The *River GeoDSS* implements a single link to model return flows to each of the links flagged with return flows; therefore, both return flows and concentration need to be combined at the

return flow link to the joining node. The flows will be added together but the resulting concentration needs to be recalculated to reflect the tributary and main stream water blend. Assuming complete mixing of the individual return flows before it reaches the node, the mass balance equation can be written as:

$$V_R C_R + V_T C_T = V_c C_c \tag{IV-11}$$

where, V_R = the volume and C_R = concentration returned to the upstream river link. V_T = the volume and C_T = concentration returned to the upstream tributary link. The combined volume and concentration are V_c and C_c. The combined volume is the sum of the incoming volumes ($V_R + V_T = V_c$). The combined concentration is computed as:

$$C_c = \frac{V_R C_R + V_T C_T}{V_c} = \frac{(V_c - V_T) C_R + V_T C_T}{V_c} \tag{IV-12}$$

Water Quality Calibration in Calibration Structures

The *flow-through* demand nodes used in the calibration structures provide water downstream with the same concentration that the water that reached the demand node. This requires special handling because the artificial connection is neglected in the network tracing algorithm. The concentration in the artificial link (*infLink*), which returns the measured water to the system as result of the *flow-through* operation, is set while calculating concentrations at the demand node. The algorithm relies on the *flow-through* demand node concentration calculation being triggered earlier than the calculation at its corresponding return node (*flow-through* Node). This is guaranteed from the network tracing calculation order. In some cases, *infLink* might contain additional flow from local inflows or other *flow-through* demands. The concentration handling at the *flow-through*

node is based on calculating an equivalent concentration using the different sources and their concentrations. The procedure calculates the mass corresponding to the local inflow and loops through the *flow-through* demands associated with the non-storage node adding the mass contributions from these nodes. This procedure is performed once at a time for all the *flow-through* demands during the upstream to downstream water quality routing.

Calibration in Simulation Runs

The simulation mode with calibration results uses the water quality calibrated concentrations to set unknown concentrations in the system. In simulation, concentration is set at: (1) inflows with no measured concentrations, (2) the calibration links (suffixed with *CALIB_SOURCE* and *_CALIB_DS_SUPPLY*) and (3) reservoir inflows. The reservoir inflow link is artificial in MODSIM; therefore, it is not available in the MODSIM output file. The concentration in the reservoir node output table is used to store and later retrieve the concentration for its artificial inflow link. This calibration concentration is different than the node concentration; therefore, the output should be interpreted accordingly.